PURSUIT OF METRIC:

OPERATIONS IN 3D RELIEF SPACE ENABLE PERCEPTION OF METRIC SLANT

Xiaoye Wang

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Psychological and Brain Sciences and the Cognitive Science Program
Indiana University
May 2019

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Doctoral Committee

_____
Geoffrey P. Bingham, Ph.D.

_____
Thomas A. Busey, Ph.D.

_____
Jason M. Gold, Ph.D.

_____
Franco Pestilli, Ph.D.

Date of Defense – April 5, 2019

Acknowledgements

A lot of things could happen in six years. A lot of things did happen in the past six years that I spent here in Geoff's lab at IU. The most important of them all was not to have finished my degree, but rather, was all those mistakes that I have made and challenges that I have faced along the way. Although I do wish that I had not made some of the mistakes and that some of the challenges were easier to handle, it was those mistakes and challenges that have eventually led me to where I am now. Of course, I cannot possibly solve those problems all on my own and it was the help that I received along the way that I am most grateful of.

感谢我的家人，尤其是我的父母。离家十年，每次回去都能注意到你们的发鬓会多出一些白发，眼角的皱纹也变多了，但是不变的是你们给我带来的家的感觉。我这些年犯了很多错误，也收到了不少打击，但是你们都在我身边，和我一起度过这些难关。我唯一懊悔的是没办法和你们一起度过更多时光。谢谢你们给予我的培养、信任和支持。没有你们我是无法得到现在我所拥有的任何东西。

Thank you, Harry, you brought me into the world ecological psychology. The deep conversations we had on the philosophical foundation of ecological psychology laid the groundwork of what I am doing now. The ideas that we exchanged captivated me and were the reason that I decided to pursue this degree. You showed me the greater philosophical context in which everything that I empirically investigate lays. What you have taught me has become an instinct that guides my analysis of the questions that I examine.

Geoff, I cannot say thank you enough for uplifting me to where I am now. I still remember the evening on my grad school visit when we were sitting by the fire in your house and sharing ideas about research. Frankly, when I first came to the lab, I had absolutely no idea of what was going on. The research and experiments that took place in this lab were all seemed

to be extremely technical, and even esoteric. However, with your patience and persistence, I gradually picked up the pace and now that I look back at everything, it is truly amazing how far I have gone. So, thank you for being such a great teacher. In the meantime, aside from being my mentor on research, you have also become a friend, a father-like figure even. I have been through a lot of things during my years at IU and you bear witness to all of them. The opposite is true, as well. You offered me guidance during the difficult times that I had, pushed me when I felt like giving up, and righted me when I deviated from the path. Thank you for believing in me. I will miss you dearly.

Mats, thank you for being a great teacher. Your knowledge on structure-from-motion and 3D perception has helped me to go through the toughest time during my research. Although we have been primarily working through Skype/Facetime and email, I do appreciate having the privilege of having you here in Bloomington and work with you face-to-face in the spring of 2019. Our discussions were enlightening and generated a lot of interesting ideas for the future. It was with your help, I finally started to appreciate what you mean by *motion* and structure from motion. I have evolved from someone who does not even understand the meaning of affine to who is relatively comfortable with all the modeling and analysis. I cannot do this without your help. So, thank you.

For all of those who were and are in the perception and action lab, Sam, Winona, Aaron, Olivia, and Rachel, I thank you for being there with me during my journey at IU. It was really fun to work with all of you and you have certainly taught me a lot. I also quite enjoy your presence at the lab gathering and the fun and difficult times that we had together.

I am really reluctant to write this acknowledgement. Finishing this only means that my graduate school career at IU has finished. For all of the good times and bad times, for all the

excitement of a significant result and all the loathing for failures to confirm my hypotheses, I thank all of you for being there with me. Although I had been constantly thinking about leaving Bloomington, go somewhere that is more fun and has more friendly weather, overtime, I have gotten to be quite fond of it and do not wish to leave. Life goes on, as people often say. Yes. Life indeed goes on and it is the things that have passed made me who I am. For that, I thank all of you for being there with me.

Xiaoye Wang

PURSUIT OF METRIC:

OPERATIONS IN 3D RELIEF SPACE ENABLE PERCEPTION OF METRIC SLANT

Perception of three-dimensional (3D) objects has been characterized in terms of relief structure, meaning the object's frontoparallel dimension is perceived accurately while its depth is scaled by an unknown factor. Previous research has identified a bootstrap process, in which observers identify and track a right angle on an object through large continuous perspective change to recover the object's width-to-depth aspect ratio accurately. In this dissertation, I aimed to apply the bootstrap process to account for 3D metric slant perception. In a series of six experiments, I made four claims. First, the bootstrap process requires relief structure. This was tested in the first and second experiments by using either strictly planar surfaces that lacked a fourth non-coplanar point required to produce relief structure or instead, surfaces with the requisite non-coplanar points to yield relief structure. Second, the bootstrap process does not require trackable 2D optical points but instead, uses trackable 3D relief structure, as shown in the second experiment where dynamic random dot displays yielded similar performance as those with trackable points. Third, the original right-angle solution is not suitable to account for recovering metric slant, as tested in the third experiment where symmetric hexagonal surfaces were used to eliminate right angles in the object. Fourth, symmetry aides the recovery of slant but perspective change aids the recovery of the scaling factor, as tested in the fourth experiment where asymmetric pentagonal surfaces were used and the fifth experiment where the hexagonal object's symmetry axis was perturbed. Finally, the last experiment replicated these findings using 3D polyhedrons. I proposed an alternative solution to the bootstrap process that only

requires identifying and tracking two equidistant points. I presented simulation results based on this new solution, which further confirmed the claims that were made.

_____
Geoffrey P. Bingham, Ph.D.

_____
Thomas A. Busey, Ph.D.

_____
Jason M. Gold, Ph.D.

_____
Franco Pestilli, Ph.D.

Table of Contents

Chapter I: General Introduction

We live in a three-dimensional (3D) world. Psychologists and cognitive scientists alike have been investigating extensively how we perceive such a 3D world. The rationale behind such an investigation is that our retina could only form 2D images based on projection, rendering the loss of the depth dimension that needs to be recovered for essentially every daily activity. This view, perhaps, could be most readily understood from Bishop Berkeley's claim in his *An Essay Towards a New Theory of Vision* (1709/2002):

> It is, I think, agreed by all that Distance of it self, and immediately cannot be seen. For Distance being a Line directed end-wise to the Eye, it projects only one Point in the Fund of the Eye. Which Point remains invariably the same, whether the Distance be longer or shorter (sect. 2).

In this passage, the phrase "[a] Line directed end-wise to the Eye" has become the epicenter for various ensuing perceptual and cognitive theories that aimed to address different perceptual phenomena. One area that attracts a lot of attention over the years is 3D shape perception.

A common approach in studying 3D shape perception is to examine the mapping of physical space (Φ) onto psychological or perceived space (Ψ). This mapping could be described using certain mathematical relations, where the perceived space is a function of the physical space, $\Psi = f(\Phi)$. These possible mathematical relations are called the *geometry of perceptual space*, an idea borrowed from Felix Klein's Erlangen Program (Klein, 1893). Based on Klein's conception, different types of geometries are defined by the types of *invariant structures* after certain *structural transformations* (Klein, 1893). The structural transformations in Klein's theory correspond to the mapping between the physical and the perceptual space, whereas the invariant structures in the context of 3D shape perception are the properties of an object that could be perceived accurately by the perceiver, since these properties would remain the same in both physical and perceptual space (e.g. Tittle et al., 1995; Todd, 2004; Todd et al., 2001; Wagner,

2008). On the other hand, properties of an object that do not remain invariant over the transformation would be perceived with some systematic errors.

In essence, there are two kinds of geometries that are commonly invoked while empirically examining 3D shape perception: *Euclidean*/*metric* and *affine* geometry (e.g. Todd, 2004). For *Euclidean* or *metric geometry*, the distance between any pair of points on the object will be preserved after any arbitrary rigid translation and rotation. Based on Klein's framework, the type of transformations associated with Euclidean geometry is rigid translation and rotation, whereas the invariant structure includes the distances between points and any arbitrary angles on the object. The implication of this kind of invariant structure is that properties such as the absolute length of any line segment, angles formed between any two line-segments, and the overall shape are preserved after the transformation. With respect to the geometry of perceptual space, metric geometry is equivalent to a one-to-one mapping between physical and perceived space, yielding an accurate representation of the 3D object. Thus, if the perceptual space is best described by Euclidean geometry, then one would expect performance is accurate in tasks such as judging the absolute distance between two points, bisecting angles between two line segments on the object, and judging object shapes (Wagner, 1985).

The second type of geometry is called *affine geometry*. Formally, affine geometry preserves properties after transformation by parallel projection (Birkhoff & Mac Lane, 1996). These properties include collinearity of points, parallelism between line segments, and ratios of lengths of line segments. In the context of the geometry of perceptual space, however, affine geometry, or affine structure has been used widely to refer the homogeneous stretching or compression of distances orthogonal to the horizontal and vertical axes, i.e. the frontoparallel plane (e.g. Domini & Caudek, 2013; Koenderink & van Doorn, 1991; Todd & Bressan, 1990;

Todd & Norman, 1991; Todd et al., 2001; Wagner, 1985). Having lengths or extents being scaled by an unknown scaling factor along the depth dimension while being metrically specified in the projection plane allows more operations than those defined by an affine mapping only. Therefore, to avoid confusion, I shall refer to this perceptual phenomenon as *relief structure*, which is in accordance with Koenderink and van Doorn (1991) as well as Lind (1996).

The implication of relief structure in the context of human perception is that ratios of distances between parallel line segments are preserved, while those between nonparallel line segments are systematically distorted. In other words, only the relationships among nonparallel lines are affected. Applying this idea to the geometry of perceived space, relief structures produce a perceived 3D object that is scaled by an unknown scaling factor parallel to the line of sight but has an accurate representation of the object's absolute sizes in a frontoparallel plane, which is perpendicular to the line of sight. Therefore, if the mapping between the physical and perceptual space is best described by the relief structure, then one would expect accurate performances in tasks such as distance bisection in any given direction, computing metric length ratio of two parallel line segments, producing a series of nominal categorizations, such as differentiating between planar and non-planar configurations, and determining the structural differences between any pair of objects that cannot be made congruent by an affine stretching transformation in depth (e.g. Lappin & Fuqua, 1983; Purdy & Gibson, 1955; Todd & Bressan, 1990; Ullman, 1983). However, one should also expect systematic distortions in performances on tasks such as judging the relative lengths and angles between two nonparallel lines on an object and bisecting an angle on the object (Wagner, 1985; Norman, Todd, Perotti, & Tittle, 1996).

Numerous studies in 3D shape perception have shown that perceived shape is typically compressed or expanded in depth (e.g. Todd et al., 1995; Todd et al., 2001). Specifically, the recovery of 3D structure has been found to be inaccurate with monocular structure-from-motion (SFM; Norman & Lappin, 1992; Norman & Todd, 1993; Perotti, Todd, Lappin, & Phillips, 1998; Tittle et al., 1995; Todd & Bressan, 1990; Todd & Norman, 1991), with stereopsis (Johnston, 1991; Tittle et al., 1995), with the combination of SFM and stereopsis (Tittle & Braunstein, 1993; Tittle et al., 1995), and with the combination of various other visual information (Norman & Todd, 1996; Norman, Todd, & Phillips, 1995). Such evidence suggests that 3D shape perception should be inherently of relief structure. However, if this were the case, daily tasks such as reaches-to-grasp objects would be seriously impaired, as they require accurate perception of metric shape to scale the grasp formed on approach to an object, which is most typically grasped back-to-front more than side-to-side. This was shown by Lee, Crabtree, Norman, and Bingham (2008) to be the case for feed-forward reaches-to-grasp, where the authors argued that poor shape perception resulted in poor reaches-to-grasp when the hand was not visible during the reach and became accurate when it was, allowing the use of online guidance. Ultimately, the fact that most studies report poor 3D shape perception is at odds with the common functionally effective and efficient performance of reaches-to-grasp, whether the hand is visible or not. Indeed, Lee and Bingham (2010) subsequently revealed conditions that yielded good levels of performance.

1.1 From Relief to Metric

Bingham and Lind and colleagues suggested that the discrepancy might be the result of experimental design because most studies that report such results presented their stimuli either

statically or with a small amount of relative motion, while people in their daily activities are moving substantially relative to their surroundings (Lee & Bingham, 2010; Lee, Lind, Bingham, & Bingham, 2012; Lind et al., 2014). To address this issue, their initial idea was as follows: because distance can be perceived accurately in a frontoparallel plane, but not in the depth direction, one can simply exchange depth and width by rotating the object, or moving the observer relative to the object, by 90°, so that the depth, now placed in a frontoparallel plane, can be perceived accurately. Upon investigation, this worked! They then tested a series of smaller amounts of relative motion and found that with continuous perspective change of approximately 45° or greater, people could perceive metric depth, and thus shape, accurately.

Lind et al. (2014) proposed a bootstrap process to account for this phenomenon. This process first assumes that observers were able to recover 3D relief structure through two-frame apparent motion, which a number of existing structure-from-motion (SFM) models can achieve (e.g. Koenderink & van Doorn, 1991; Shapiro, Zisserman, & Brady, 1995; Lind, 1996). The resulting depth structure is the actual depth scaled by an unknown factor. As shown in Figure 1, to recover the correct scaling factor, one first identifies two points on the object that are equidistant to the observer (A_0 and B_0), an operation allowed in the relief frame of reference since depth is scaled uniformly by the unknown scaling factor. With the assumption of small visual angle and, therefore, scaled orthographical projection, the line formed between these two points is orthogonal to the line of sight. Consequently, a third point (C_0) can be identified on the object that is on a line parallel to the line of sight, i.e. $\angle C_0A_0B_0 = 90°$. These three points form a right angle that can be tracked across perspective change. Assuming rigidity, the angle formed among these three points ($\angle C_1A_1B_1$ of Figure 1, right) would remain as 90° on the object but would be constantly changing in the *relief space* as a result of the interaction between motion in

depth and the relief scaling factor. Subsequently, one can relate the new angle to the original 90° angle to recover the scaling factor. When the perspective change is small, the new angle typically fails to deviate from the original angle by a sufficient amount to discriminate such change due to motion measurement noise. However, with sufficiently large perspective change, variations in the target angle in relief space would be large and therefore, produce relatively more stable and accurate estimation of the scaling factor. In addition, the unknown scaling factor renders the amount of rotation as unknown, and consequently, observers cannot know when perspective change is sufficiently large to "trust" the resulting scaling factor estimate. Therefore, there has to be something to inform them when this has occurred. Because angle bisection is allowed in relief space, the bisection of the target angle by the line of sight can be detected to inform observers that a 45° perspective change has occurred, and the recovered scaling factor should be accurate and reliable, i.e. when $\angle C_1A_1D = \angle DA_1B_1$.

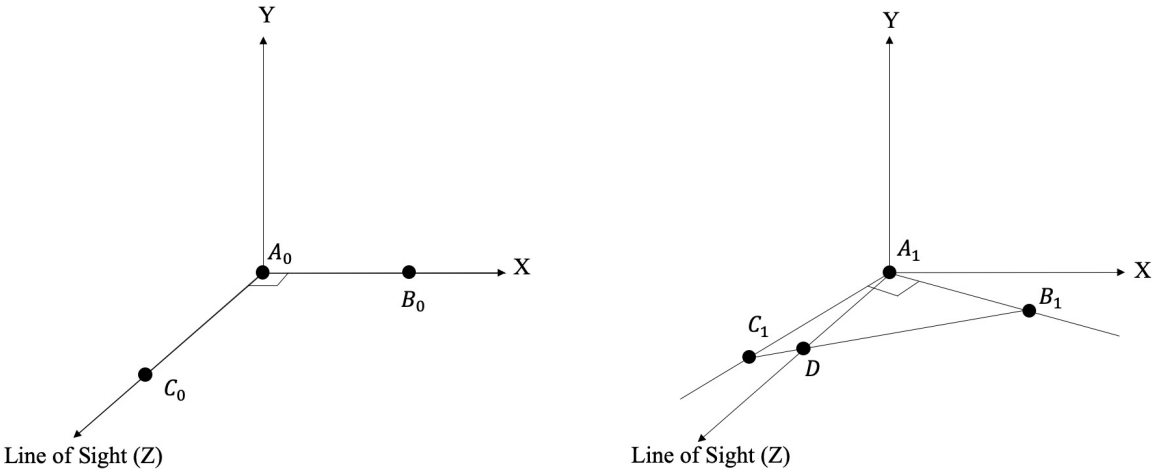


Figure 1. Illustration of the original right-angle solution to the bootstrap process. At t_0 (left), two points that are equidistant to the observer (A_0 and B_0) can be first identified and a third point (C_0) can be found on the object that is on the line of sight or a line that is parallel to the line of sight. By definition, these three points for a right angle, i.e. $\angle C_0A_0B_0 = 90°$. As the object

rotates, at t_1, $\angle C_1A_1B_1$ is no longer 90° in the relief space due to the unknown scaling along the line of sight (the Z axis). Assuming rigidity, the correct scaling factor would be the one that could make $\angle C_1A_1B_1 = 90°$. At rotation less than 45°, the line of sight fails to yet bisect the angle (shown on the right). The resulting scaling factor is the most reliable when the line of sight bisects $\angle C_1A_1B_1$, i.e. when $\angle C_1A_1D = \angle DA_1B_1$.

The experimental results consistent with the bootstrap process have been replicated in a number of different studies and tasks, including reaches-to-grasp wooden elliptical cylinders varying in depth-to-width aspect ratio (Lee & Bingham, 2010), recognition of complex computer generated polyhedrons evaluated in terms of both errors and reaction times (Lee, Lind, Bingham, & Bingham, 2012), and judgments of 3D depth-to-width aspect ratios of computer-generated 3D elliptical cylinders and asymmetric polyhedrons (where judgments were performed by adjusting the aspect ratio of an outline of the shape) (Lind et al., 2014).

What one needs to note is that the essence of the bootstrap process is not necessarily 45° of continuous perspective change or bisecting an arbitrary right angle on the object; instead, it is identifying and consequently tracking an *invariant over the SFM transformation*. In the context of bootstrapping metric depth via relief depth, the invariant corresponds to identifiable metric or rigid physical structure of the object, while the transformation corresponds to the constant changes in the perceived relief structure as a result of relative motion in depth due to continuous perspective change. For the bootstrap process to work, one has to identify certain depth structures that are equivalent in both the relief and metric space at some instant of time. Such an equivalence offers a bridge that links relief space to metric space. In the solution presented by Lind et al. (2014), the right-angle acts as such a bridge. Because depth is scaled in relief space by

the same scaling factor, albeit unknown, identifying two equidistant points is allowed. Also, assuming small visual angle and therefore scaled orthographical projection, a third point can also be found such that the three points form a right angle. This angle is also of 90° in the physical space and therefore should remain invariant as the object moves relative to the observer. However, because relative motion creates depth changes, this angle would no longer be 90° in relief space. The observer, therefore, can use such transformation to determine the scaling factor that would yield the angle as 90° again.

1.2 Before Shape, There Is Slant

3D slant perception has been treated as a component of 3D shape perception (e.g. Koffka, 1935; Gibson, 1950; Beck & Gibson, 1952; Wallach & Moore, 1962; Kaiser, 1967; Hoffman & Richards, 1984; Todd, 2004; Sakata, Tsutsui & Taira, 2005; Welchman, Deubelius, Conrad, Bültoff & Kourtzi, 2005). Most notably, Gibson (1950) argued that the perception of a 3D shape should be an amalgam of perceived distance and slant at every local region (although Gibson did later renounce such argument in his seminal work *The Ecological Approach to Visual Perception* (Gibson, 1986)). Such a claim was subsequently adopted by Marr in his famous 2.5D Sketch, which utilizes viewer-centered representations of local features to extrapolate 3D objects based on 2D retinal images (Marr, 1982). Under this view, a 3D polyhedral shape is composed of a set of local surfaces each of which would be perceived in respect to its 3D orientation. In turn, the perceived 3D orientation of a planar surface can be specified via its local surface orientation. As shown in Figure 2, surface orientation can be decomposed into three components: slant (σ), tilt (τ), and roll (ω). Traditionally, slant has been defined as the angle formed between the line of sight, which corresponds to the z-axis, and surface normal ($\vec{n}$). Tilt is the angle between the

projection of the surface normal on the image plane ($\vec{n}_{xy}$) and the x-axis, indicating the direction of slant (i.e. the direction of the maximum increase in distance relative to the line of sight) (Stevens, 1983a, 1983b). Roll is simply the angle of rotation of the surface around its normal. For a planar surface, roll is rather hard to define due to a lack of reference direction and is not as commonly studied as slant and tilt. Defining surface orientation this way can be found in many different studies, including those of 3D shape perception (e.g. Gibson, 1950; Gibson & Cornsweet, 1952; Stevens, 1983a, 1983b; Saunders & Knill, 2001; Norman et al., 2006; Sawada & Pizlo, 2008).

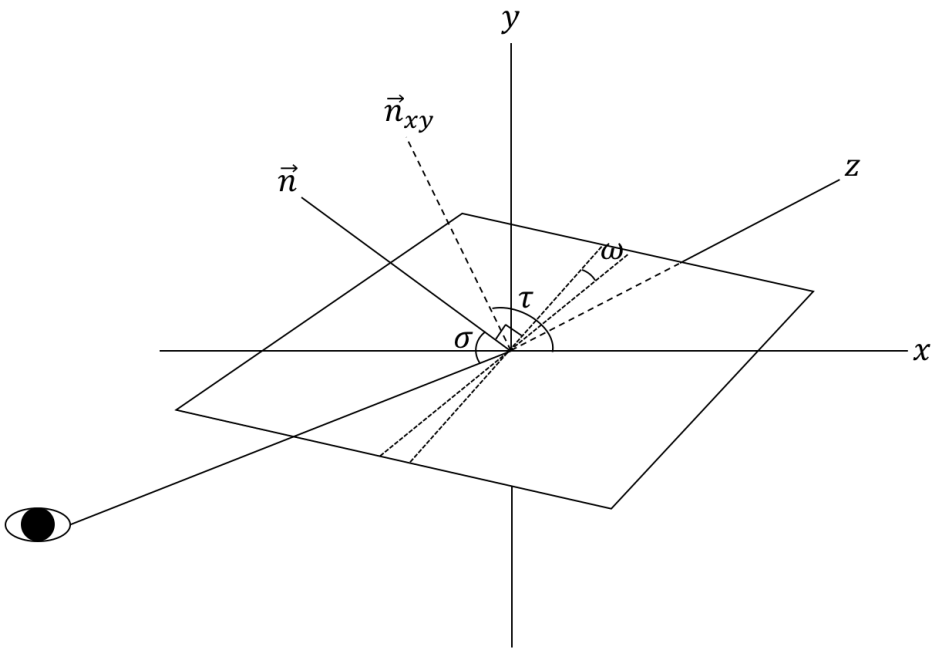


Figure 2. Illustration of variables that are used to define surface orientation. Slant (σ) is defined as the angle formed between surface normal $\vec{n}$ and the line of sight. Tilt (τ) is defined as the angle between the projection of the surface normal on the xy-plane $\vec{n}_{xy}$ and the x-axis. Roll (ω) is the angle of rotation of the surface round its normal.

The fundamental assumption in this understanding of 3D shape perception is that local perceptual estimates of 3D surface orientations are composed to yield perceptions of global 3D object shapes. This approach has been elaborated to address the perception of 3D smoothly curved shapes to entail perceptual estimates of local higher order shape specific surface properties (e.g. Perotti, Todd, Norman & Phillips, 1998). Accordingly, Todd (2004) argued that 3D shapes were represented through a local map consisting of different neighborhoods that depict different aspects of the local 3D structures. Together, these local neighborhoods specify a particular 3D shape as a whole. Thus, according to these theories, the ability to perceive a 3D shape entails the ability to perceive local 3D structures of the object. In the case of 3D polyhedral shapes, this would entail perception of local surface slant.

Similar to findings in 3D shape perception, many studies have also reported inaccuracies in slant judgments. For instance, Gibson (1950), using a palm board as a response method for judging slant, found that perceived slants tended to be biased towards the frontoparallel plane, i.e. appeared to be less depthy, when the slants were presented using texture gradients. Additionally, Flock (1964) demonstrated that such underestimation could be alleviated when the slanted surface is in lateral motion. In more recent studies, slant perception has been investigated using discrimination tasks (e.g. Knill & Saunders, 2003; Hillis, et al., 2004). In contrast, Todd, Christensen, and Guckes (2010) used a slant-matching task in which participants were asked to adjust the orientation of a line to match the 3D slant of a surface shown in a display of a static texture gradient with a relatively small field of view. Using this method, the authors found that the perceived slant angle was systematically underestimated. They also found that static binocular disparities enabled a more accurate perception of slant.

Based on the success of the bootstrap process in addressing metric 3D shape perception, it might be possible to apply the same process to yield the perception of 3D metric slant. However, the issue with adopting a local frame of reference in the experimental paradigm that has been used to study the bootstrap process is that the local slant itself would not remain constant over the SFM rotation (or perspective changes). Specifically, assuming the observer fixates on the center of the surface, as the object rotates, the angle formed between the line of sight and surface normal would be constantly changing. Figure 3 shows the local slant of a surface slanted 30° at 90° local tilt (i.e. the surface is facing the observer) as the surface rotates around a vertical axis that passes through its center. Throughout the rotation, the surface maintains a 30° angle to the ground plane. As can be seen from this figure, there is a large variation in local slant (≈10°) as the surface rotates 45° (i.e. 45° change in tilt). Therefore, using a local frame of reference to describe slant in the context of SFM is not effective and could be extremely confusing simply because the slant, if defined locally, is not constant.

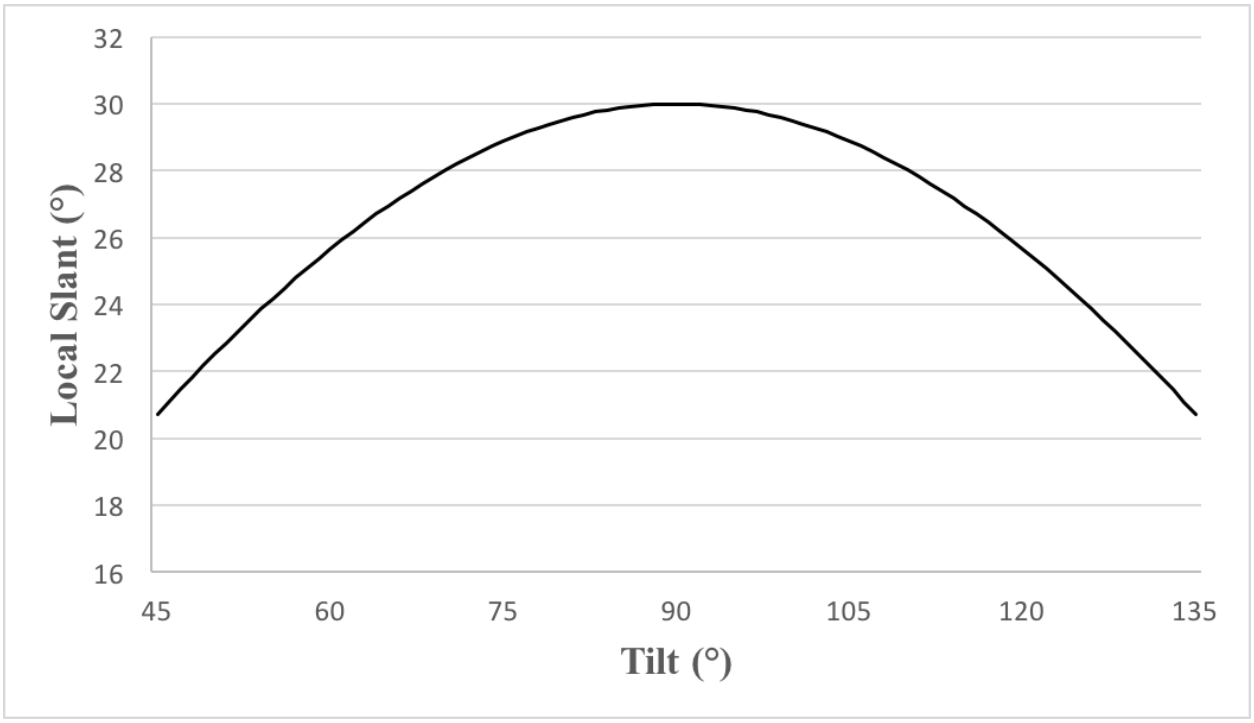


Figure 3. Local slant variation as a surface that forms a 30° angle with the ground plane rotates around a vertical axis that passes through its center. 90° is when the surface is facing the observer's line of right.

Gibson (1950) offered an alternative definition of slant that he called gravitational or geographical slant. This is defined as the angle formed relative to the horizontal and vertical axes determined by gravity, and in the current context, the angle formed between the object and the ground plane (p. 369; see also Gibson and Cornsweet (1952), and Sedgwick and Levy (1985)). Geographical slant is defined relative to the environment and remains constant over rotation around a vertical axis. Mathematically, this angle can be found through computing the angle between the object's surface normal and the ground plane's normal, which can be described using the unit vector pointing towards the y direction [0, 1, 0]. Correspondingly, tilt in this frame of reference should be defined as the projected surface normal on the xz-plane and the unit vector pointing towards the z direction [0,0,1]. Figure 4 shows this. I would refer to these as *geographical slan*t and *tilt* because slant remains invariant as the line of sight changes and tilt describes the direction of the gradient of slant relative to the ground surface. Based on this description, a geographical slant of 90° would be an upright surface and a slant of 0° would be horizontal (parallel to the ground). As the slanted surface rotates around a vertical rotational axis, which coincides with the y-axis, its geographical slant would remain constant while its local slant would be constantly changing.

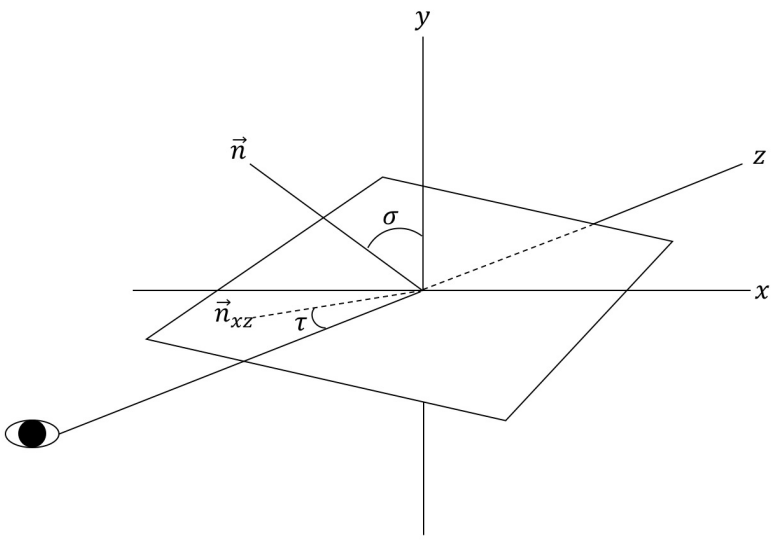


Figure 4. Illustration of geographical variables in the current context. Geographical slant (σ) is defined as the angle formed between surface normal $\vec{n}$ and the positive y-axis. Geographical tilt (τ) is defined as the angle between the projection of the surface normal on the xz-plane $\vec{n}_{xz}$ and the z-axis.

Logically, if 3D shape perception can be improved with large continuous perspective change, and if slant perception is a component of the process yielding 3D shape perception (for polyhedral objects), then we would expect performance in slant perception to improve to become veridical with sufficiently large continuous perspective change. The goal of this dissertation is to explore the applicability of the bootstrap process to 3D slant perception. The first issue that may arise when applying the bootstrap process to 3D slant perception is related to the recovery of relief structure. The bootstrap process operates in 3D space and thus assumes the available optical information could enable the perception of relief structure. The model then applies relief operations with sufficiently large perspective changes to obtain metric or Euclidean structure. The problem is that slant perception usually entails displays of planar surfaces. The SFM models, on the other hand, require four non-coplanar points (Koenderink & van Doorn, 1991; Lind, 1996;

Shapiro, Zisserman & Brady, 1995). 3D shapes under rigid rotation in SFM displays would entail non-coplanar points, and thus, yield relief structure, but planar surfaces rigidly rotating in monocular SFM displays would not.

Additionally, another issue that may emerge regards the direction of slant, i.e. tilt. In a local frame of reference, the direction of slant is simply the direction in which there is a maximum increase in *depth* relative to the observer's line of sight on the surface (e.g. Stevens 1983a, 1983b). In a geographical frame of reference (i.e. for geographical slant), this property can no longer be used to identify the direction of slant, since tilt is now defined in terms of the direction of maximum increase in *distance* relative to the ground surface. To be able to accurately perceive geographical slant, the direction of slant also has to be accurately perceived. A potential source of information that informs the observer of the direction of geographical slant is symmetry. According to Pizlo and colleagues, 3D symmetry could be a strong a priori constraint on 3D shape perception (Pizlo, 2010; Pizlo et al., 2010; Li et al., 2011). Similarly, Saunders and Knill (2001) showed that symmetry also plays a significant role in slant perception. For a symmetrical planar surface whose symmetry is described using a vertical symmetry axis and symmetry lines that connect symmetrical points around the symmetry axis, the angle between the symmetry axis and lines after projection can constrain the interpretation of slant and tilt of the surface (see Figure 3 of Saunders and Knill (2001)). Given that numerous studies have shown that observers readily detect the symmetry axes in a 2D shape oriented in a 3D environment (e.g. Bingham & Muchisky, 1993a, 1993b; Palmer, 1985; Pizlo et al., 2010; Sawada & Pizlo, 2008), it is possible that the symmetry axis could provide information about the direction of slant in a geographical frame of reference.

Finally, all of the extant studies providing results supporting the bootstrap process have involved stereovision together with the continuous perspective changes in SFM displays (e.g. Lind et al., 2014). When objects move with stereo viewing, three different types of motion generate information including two forms of stereomotion in addition to monocular SFM. The stereomotion information is change-of-disparity-over-time (CDOT) and inter-ocular velocity difference (IOVD; Cumming & Parker, 1994; Shioiri, Saisho, & Yaguchi, 2000; Nefs, O'Hare, & Harris, 2010; Allen, et al., 2015). CDOT consists of the temporal rate of change of stereo disparity as the object and perceiver move relative to one another. IOVD consists of stereo disparities between the monocular optic flows in each eye. Although Lind et al. (2014) used the combination of all three types of visual information in their study, their model only required optical information that would yield relief 3D object structure then used for the bootstrap to Euclidean structure. Either monocular SFM or static stereo disparities should in principle be sufficient to yield relief 3D shape, but not the slant of a planar surface. On the other hand, stereomotion is above 1st order because it entails more than the mere two views assumed to be available for SFM or stereo. Stereo disparities entail two views, but the disparities themselves transform over time in stereomotion yielding effectively more than two views. Stereomotion information in addition to monocular optic flow for SFM may enable perception of relief structure even with planar surfaces. It remains unclear whether large perspective changes would be effective in perception of the slant of planar surfaces with only stereomotion (either CDOT or IOVD), or with the combination of stereomotion and monocular SFM. It is unlikely to be effective with only monocular SFM information. In addition, there are more issues related to 3D slant perception when considering it in conjunction with the original right-angle bisection method of the bootstrap process, which shall be discussed in the subsequent section.

1.3 Caveats of the Right-Angle Solution

Remember, the role of large continuous perspective change in the bootstrap process is to produce sufficiently large depth variations in the relief space on the invariant depth structure, e.g. the right angle, that would subsequently allow perceivers to bootstrap to metric. This means that there needs to be sufficiently large depth variation in the right angle. Originally, Lind et al. (2014) worked with aspect ratio judgments, using polyhedral objects with a horizontal top surface. Although they varied the optic slant of the top surface, defined as the angle formed between the line of sight and the surface normal, by changing the relative distance between the viewers' eyes and the object's top surface, the optical slant itself remained relatively large (either 82º or 74º). This enabled one to perform the right-angle solution because large continuous perspective change could be sufficiently translated to the depth variations of the right angle on top of the object, allowing bisection with the line of sight. However, with a more upright surface, the optical slant would become smaller (for a frontoparallel surface, the optic slant is 0º, i.e. the line of sight coincides with the surface normal). Under such circumstances, relative motion between the observer and the object cannot translate to angle variations in relief space and the angle itself would remain at around 90º.

For instance, let us compare two slanted surfaces, one that is more horizontal with a surface normal that is 10° from the y-axis (i.e. a geographical slant of 10°) and another that is more upright that has an angle of 80°. Following my previous definition of geographical tilt, a tilt of 0° would be when the surface is facing the observer. Under this setup, it is feasible to find two equidistant points on the two surfaces along with a third point that is on a line parallel to the line of sight. Rotating the surfaces back and forth around the y-axis could produce the desired

perspective change. A rotation of 65°, adhering to the previous definition of tilt, corresponds to a tilt variation between -32.5° and 32.5°. Figure 5 shows the right angles on two different slants in relief space as this happens. As one can see, when the surface is flatter, the variation of angle in relief space is large, compared to when the surface is more upright. Although the simulated tilt variation went beyond the originally identified 45° of perspective change, the angle variation for the upright surface is still small and should be incapable of producing reliable estimates of the relief scaling factor. This indicates that if the right-angle solution was used to recover metric slant, with large continuous perspective change, only flatter slants would be perceived accurately while the perception of more upright slants would still remain inaccurate. If alternatively, performance was veridical across different slants for a large amount of continuous perspective change, then the of proposing an alternative implementation of the bootstrap process would emerge.

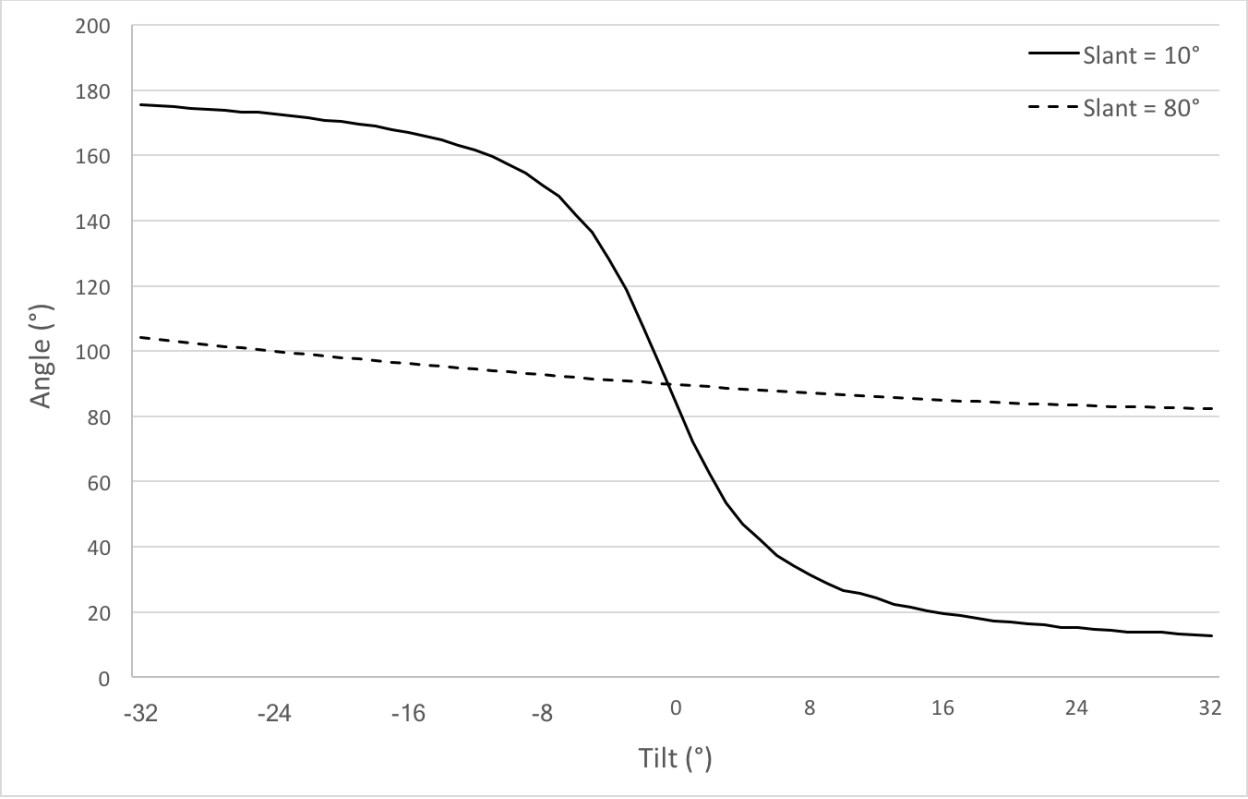


Figure 5. The angle variations in relief space for the right angle identified when tilt was 0° for a surface with a slant of 10° (flatter) and one of 80° (more upright) as a function of perspective change.

Furthermore, the original right-angle solution of the bootstrap process requires observers to identify three points on the object that form a right angle. If the display only contains CDOT, implemented through a dynamic random dot stereogram (Julesz, 1971), then there would be a lack of trackable texture elements with which an observer could formulate right angles. Therefore, if the original right-angle solution operates on the basis of trackable texture elements, then performance should not improve with CDOT-only displays. However, if performance did improve, this suggests that the bootstrap process, according to its fundamental mechanism, should operate on the level of relief structure.

1.4 Dissertation Outline

In this dissertation, I investigated issues of bootstrapping metric structure in the context of 3D slant perception, combined with an exploration of the roles of different types of visual information: in what way could large continuous perspective change enable accurate slant perception? In Experiment 1, I used rectangular planar surfaces, presented using CDOT, monocular SFM, or the two combined (which also gave rise to IOVD), and tested whether 3D slant perception could become accurate with large continuous perspective changes. Results from Experiment 1 showed that performance in the two stereo conditions (CDOT only and the combined condition) were different from the monocular SFM condition, but none of the conditions produced veridical performance when given large continuous perspective change.

Such results were somehow expected since the lack of 3D structure (i.e. the fourth non-coplanar point) of the stimuli could prevent one from deriving relief structure in the first place.

In Experiment 2, I introduced nine cuboids, laying in a rectangular grid, in addition to the rectangle planar surfaces used in Experiment 1. Performance improved in all three visual information conditions, where performance was poor at small rotations, it became veridical at large rotations and remained relatively stable with further increases of rotation. The significance of Experiment 2 was twofold. First, it demonstrated that the bootstrap process could be applied to 3D slant perception. This finding, to some extent, suggested that the original right-angle solution to the bootstrap process might not be suitable to account for 3D slant perception, since this solution implied that judgment should remain poor for more upright slanted objects. Second, in the pure stereomotion condition, because every consecutive frame is re-randomized to eliminate monocular SFM information, there were no trackable texture elements on the objects. Nevertheless, trackable points are required by the original right angle solution of the bootstrap process. This indicated that perceivers could directly use 3D relief structure to perform the bootstrap process and recover the metric scaling factor. Because the objects used in this experiment contained inherent right angles (i.e. rectangular slant surface, cuboids, and layout of the cuboids), it was essential to rule out the possibility of observers using right angles to bootstrap metric slant.

In Experiment 3, I eliminated the inherent right angles in the object, using symmetric hexagonal surfaces with nine tetrahedrons at random locations, tested using CDOT and the combination of CDOT and monocular SFM. Performance still improved with large continuous perspective change for both types of visual information, replicating the pattern of results predicted by the bootstrap process. Results from this experiment, in particular the CDOT

condition where both trackable texture elements that could potentially form a right angle and inherent right angles in the object were absent, confirmed that the right-angle solution is indeed not suitable for describing 3D metric slant perception mediated by large continuous perspective change.

In Experiment 4, I explored the role of symmetry in specifying the direction of geographical slant using irregular pentagonal surfaces with nine tetrahedrons at random locations, presented using CDOT or combined information. Performance was poor and failed to improve with an increasing amount of perspective change with either of the information types. This suggested two alternative explanations. The first is that the bootstrap process requires symmetry to recover metric structure. This interpretation seems to be too limiting for the bootstrap process and does not fit results from the previous aspect ratio judgment study (i.e. Lind et al., 2014) where the authors used an asymmetric pentagonal surface as the target object's cross-section. The second interpretation is that symmetry disambiguated the direction of geographical slant. For the asymmetric pentagonal objects, even though the correct scaling factor could be recovered through the bootstrap process, the ambiguity in the direction of slant perturbed the final slant estimation.

In Experiment 5, I used the hexagonal objects from Experiment 3 to explore the role of symmetry. I changed the roll of the object so that the symmetry axis was 15° from the actual direction of slant. Slant judgments accurately reflected the surface orientation along the symmetry axis rather than in the actual direction of slant. Thus, the bootstrap process worked, but the correctly recovered scaling factor was then used to derive slant that was based on the direction of the symmetry axis.

Finally, in Experiment 6, I used generalized cone objects instead of planar surfaces with ad-hoc 3D structures. The stimuli used thus far were all rather "strange": they were planar surfaces of various shapes with strange 3D structures on top. Generalized cones, on the other hand, can be considered to be more ecologically valid. If performance in the previous experiments can be replicated in this experiment, it would increase the power of the bootstrap process. In addition, doing so could also rule out the possibility that what observers truly used during the process was the 3D structures on top of the planar surfaces. Therefore, in this experiment, I used generalized cone objects with cross sections that were rectangular, hexagonal, and an asymmetric pentagon, just like the shape of the surfaces used in Experiments 2 to 4. Performance in this experiment replicated results found in Experiments 2 to 5, where performance improved with rectangular and hexagonal generalized cone objects but not with pentagonal objects.

In light of the findings from these experiments, I then presented a stratified process for recovering metric slant. In Chapter III, I will present a detailed roadmap for implementation of the bootstrap process in the context of slant perception, starting with an SFM algorithm used to recover relief depth structure followed by the bootstrap to the correct scaling factor then used to produce slant estimation. Subsequently, I will also provide model simulation results using stimuli from Experiments 2 to 5 so I can compare simulations with empirical results. This is a test of the model itself, an evaluation of whether it is suitable for describing the achievement of accurate metric slant estimates using large continuous perspective change. The model was developed based on the results of the initial experiments and is now tested by applying it to replicate results from a number of subsequent experiments.

Chapter II: The Experiments

Experiment 1

In Experiment 1, observers judged displays of a slanted rectangular planar surface rotating around a vertical axis through its center with different amounts of rotation and different types of visual information. Displays contained one of the three types of visual information, namely, SFM, CDOT, or the combination plus IOVD. Observers adjusted the orientation of a line in the display to match the perceived 3D slant of the surface. To anticipate, I found that 3D slant perception was different from 3D shape perception to the extent that large continuous perspective change did not reproduce the findings of Lind et al. (2014). Rather, I found that performance in the stereomotion and combined conditions exhibited the same results, both exhibiting a trend for veridical judgment without actually achieving it, whereas performance in the monocular SFM condition remained consistently poor.

Methods

Participants

Thirteen adults, age between 20 and 30 (three males and ten females) participated in this experiment, all of who provided their informed consent prior to participating in the experiment with the approval from Indiana University's Institutional Review Board (IRB). They were paid $10/hour. All participants had normal or corrected-to-normal eyesight, and also passed a stereo fly test (Stereo Optical Co., Inc.) that measured their stereo acuity. Participants had to be able to identify the target circle indexed by a disparity of 80 seconds of arc to be included.

Stimuli and Apparatus

Stimuli were presented on a Dell UltraSharp U2312HM 23-inch monitor (51 by 29 cm) with a resolution of 1920 × 1080 and a refresh rate of 60 Hz. I used MATLAB 2011b (the Mathworks Inc., Natick, MA, 2011) to generate the stimuli, which were presented with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007). From trial to trial, slant was varied randomly between 27° and 73° by 2° increments yielding 24 different slants. Again, since I was testing geographical slant, the slant angles were all defined in terms of angles formed between the surface and the horizontal plane. Because the center of the surface was placed at eye level, the equivalent optical slant when the surface tilt was 0° (i.e. facing the frontoparallel direction during rotation) would be 90° minus the geographical slant. The range of slant angles was chosen to avoid slants near either 0° or 90°. I presented a slanted surface rotating in the display about a vertical axis through the center of the surface with the rotation centered on a 0° tilt. Similar to the range of rotation used in Lind et al. (2014), there were five different rotation amounts: 25°, 35°, 45°, 55°, or 65°. This range would reveal improvement as rotation approached 45° and then continued accuracy for amounts greater than 45°. For each rotation amount (e.g. 25°), the surface rotated first to one side of 0° tilt by half of the amount (e.g. 12.5°) and then back and to the other side and back by the same amount. The surface rotated at a constant speed of 20°/s. This meant that even with 65° of rotation, the surface only rotated 32.5° away from 0° tilt. Displays continued the oscillating rotation until observers completed their judgments and ended the trial. Thus, even though a single speed of rotation would yield different durations for the different rotation amplitudes, this was independent of trial duration and the duration during which information was available to the observer. Todd and Bressan (1990) (and subsequently, other researchers) similarly controlled for their variations in frame number in SFM displays by continuously oscillating the display until judgments were completed.

The 24 different slants were tested using each of the 5 different rotation amounts. I also varied the length of the slanted surface while fixing the width at 10 cm. Three surface lengths (8 cm, 10 cm, and 12 cm) yielded 3 projected heights at each slant.

I used red-blue random dot anaglyphs to generate displays that contained three types of visual information: monocular motion, stereomotion, and the two combined. For all conditions, the dots had the size of one pixel, with a total of 6000 dots. The displays contained a background, placed 18 cm behind the screen, and a target surface, placed 9 cm behind the screen. The point of observation used to generate the stimuli was placed 76.2 cm in front of the screen, mimicking the actual distance during the experiment. Given the above layout, the background surface spanned approximately 9.10° horizontally and vertically. The horizontal size of the projected slant surface is 6.72°, whereas its vertical size varied as a function of slant angle and surface length, approximately ranging from 2.10° (slant angle 23° and surface length 8 cm) to 7.71° (slant angle 73° and surface length 12 cm). The size of the stimuli therefore enabled approximation of perspective projection using scaled orthographic projection, which is a crucial requirement for the model. Figure 6 is a schematic of this setup.

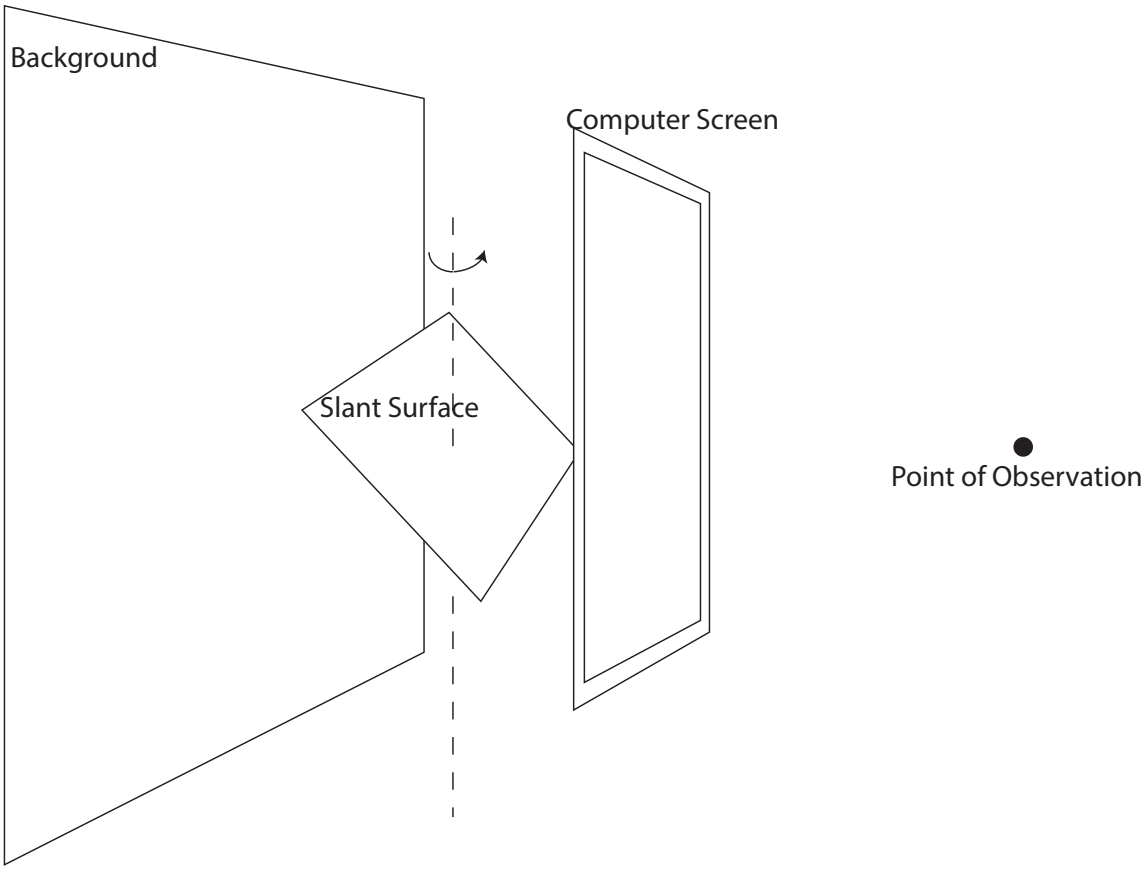


Figure 6. Schematic demonstration of the stimuli setup in Experiment 1. Note: display setup is not up to scale.

The displays were generated by first constructing the actual target surface and a background using random dots in a 3D space, and back-projecting the dots onto the screen surface through the point of observation (projection point). The back-projected dots were the ones to be drawn on the screen. Hidden line removal was used to address the occlusion of the background as the slanted target surface rotated. For both stereomotion and combined conditions, I projected the points through left and right projection points, corresponding to the viewers' left and right eyes, with an inter-pupillary distance (IPD) of 6 cm. The stereomotion condition only contained CDOT information. To do this, I re-randomized the dot positions every frame while

preserving the evolving disparities, so that only binocular disparity could specify the changing 3D structures in the display. The combined condition contained CDOT, IOVD, and monocular SFM information. The random texture was not re-randomized each frame. CDOT was specified through the evolving binocular disparity, whereas IOVD and SFM were specified through the stable monocular structure across frames. Finally, the monocular motion condition only contained structure-from-motion information. For this condition, the texture was not re-randomized each frame and the IPD was set to 0. The color of the dots in this condition was therefore the combination of red and blue. Because of the way our displays were constructed, there was also texture gradient information in the combined and SFM displays.

Because the combinations of rotation amount and slant angle yielded 120 trials for each of the three visual information conditions, I decided not to fully cross surface heights with slant angles. Instead, I treated three consecutive slant angles as a block (e.g. 17°, 19°, and 21°), and randomly assigned one of the three slant heights to each of the slants in each consecutive block (e.g. 10 to 17°, 12 to 19°, and 8 to 21°). The surface height assignment was fixed for each combination of rotation angle and slant angle of each visual information condition (e.g. all participants would see the surface height of 10 for slant angle of 17°, rotation angle of 25, in the monocular condition).

Procedure

After participants provided informed consent and passed the stereo fly test, they sat in front of the computer screen, wearing a pair of red-blue filter glasses and viewing the display binocularly for all three visual information conditions, and were instructed how to perform the task. Participants first observed a rotating planar surface with a certain slant. After the surface

completed one cycle of rotation (i.e. the surface rotated back to the starting position at 0° tilt), a 2D response line appeared on the screen to the right of the display. Participants were instructed to use the left and right arrow keys to adjust the orientation of the line so that it matched the slant of the surface in the display (as in Todd, Christensen & Guckes (2010)). Although studies have demonstrated that perception of 2D lines tends to be biased (Dick & Hochstein, 1989; Durgin & Li, 2011), Cherry and Bingham (2018) tested the effectiveness of using a 2D line as a response figure for slant judgment, compared with adjusting an actual surface, and did not find a difference between the two methods. The surface continued to rotate back and forth throughout the trial while participants adjusted the response line and until they hit the space bar to enter the judgment once they were satisfied that the orientation of the line matched the surface slant.

Each participant completed three sessions, one for each visual information condition (120 trials). The order of visual information conditions was randomly determined. Each session was conducted at least one day apart from the previous one. Within each session, participants experienced the rotation amounts in order from 25 to 65. I did not randomize the order of rotation amounts to avoid the possibility that large perspective change would calibrate the judgments performed with smaller perspective changes. Within each rotation angle condition, the slant angles were randomly displayed.

Data Analysis

For data analysis, I first used multiple regression to examine the effects of surface height on perceived slant. I found such effects to be small, so I did not treat different surface heights as a separate factor in the subsequent analyses. The goal of subsequent analyses was to explore the effects of rotation amount and visual information conditions on perceived slant. I performed

linear regressions with the actual slant as the independent variable and the perceived slant as the dependent variable. I computed the regression slopes, intercepts and r^2 for each participant in each information and rotation amount condition. Figure 7 shows examples of such regression. For a given rotation amount, greater-than-1 regression slopes and smaller-than-0 regression intercepts correspond to an underestimation of small slant and an overestimation of large slant, suggesting an overall inaccuracy in slant judgment (Figure 7 left). On the other hand, a regression slope of 1 and intercept of 0 suggest an overall accurate slant judgment (Figure 7 right). Then, I used the resulting slopes, intercepts, and r^2 as dependent measures in repeated-measures analyses of variance (ANOVA's). In these ANOVA's, there were two within-subject factors, visual information (three levels) and rotation amount (five levels). Finally, to evaluate the actual performance, 95% within-subject confidence interval around the mean calculated for within-subject design (Cousineau, 2005; with correction by Morey, 2008) was also used for regression slopes and intercepts in each information condition.

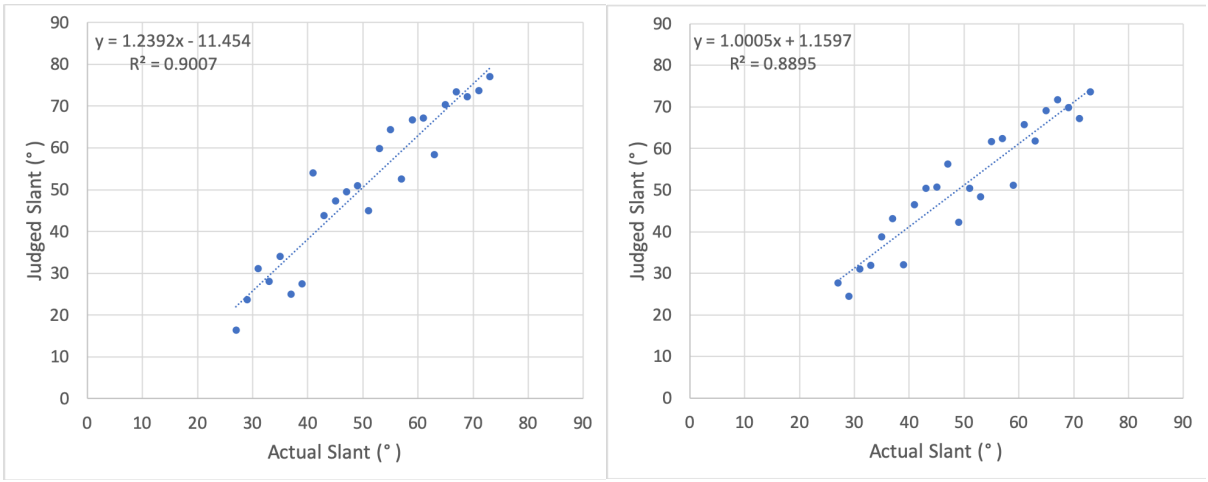


Figure 7. Example linear regression, between actual slant and judged slant for 25° (left) and 35° (right) of rotation. Performance for the 25° condition was deemed to be poor, where underestimation of small slant and overestimation of large slant led to a greater-than-1 regression slope and smaller-than-0 regression intercept. Performance in the 35° condition was veridical.

Results and Discussion

I began analysis by using multiple regression to evaluate the effects of surface height on judged slant. I evaluated the effects of three factors (rotation angle, actual slant angle, and surface height) separately within each visual information condition. The dependent variable was perceived slant. For all three visual information conditions, slant angle accounted for the majority of the variance in the model. Surface height accounted for a very small portion of the total variance in each of three visual information conditions (0.8% with Cohen's $f^2 = 0.036$ for CDOT, 2.8% with Cohen's $f^2 = 0.079$ for monocular motion, and 0.4% with $f^2 = 0.019$ for the combined condition). Because changes in r^2 with the inclusion of surface height were exceedingly small, I did not include it as a separate variable in the subsequent analyses.

Next, I performed linear regressions comparing perceived and actual slants for each unique combination of visual information and rotation amount separately for each participant. Accurate judgments would yield slopes of 1 and intercepts of 0. Good precision would yield r^2 approaching 1. Figure 8 to Figure 10 show the average regression slopes, intercepts, and r^2 and for each unique combination of conditions.

For regression slopes, ANOVA showed that there were significant main effects of visual information ($F(2,22) = 24.77$, $p < 0.001$, $\eta_p^2 = 0.69$), and of rotation amount ($F(4,44) = 7.01$, $p < 0.001$, $\eta_p^2 = 0.39$). There was also a significant interaction effect between visual information and rotation amount ($F(8,88) = 2.35$, $p < 0.05$, $\eta_p^2 = 0.18$). For the significant main effects, I used a least significant difference (LSD) adjustment for multiple comparisons as a post hoc analysis. Among the three visual information conditions, the means for the monocular condition were significantly lower than those for the stereomotion condition ($p < 0.001$), and the combined

condition ($p < 0.001$). However, there was no significant difference between the stereomotion and combined conditions ($p > 0.6$).

Because there was no difference between the stereomotion and combined conditions, I used a separate post-hoc ANOVA to compare the stereomotion and combined conditions. As expected, the ANOVA showed a significant main effect of rotation amount ($F(4,44) = 7.81$, $p < 0.001$, $\eta_p^2 = 0.42$), but no difference between the two information conditions ($p > 0.6$, $\eta_p^2 = 0.023$). LSD adjustment on different levels of rotation amount showed that the 25° condition yielded a significantly greater regression slope than the other conditions. Moreover, the regression slope in the 65° condition was also significantly lower than other conditions. Despite these changes, the regression slope failed to reach 1 at 45° or beyond. This was further confirmed by the 95% confidence intervals for both conditions, as neither visual information conditions produced regression slopes that had confidence intervals to include 1 at as much as 65° of rotation (Figure 8). A post hoc ANOVA performed on the data in the monocular motion condition yielded no effect of rotation amount ($p > 0.6$, $\eta_p^2 = 0.051$). The mean slopes were below 1, with an overall mean of 0.89 and a standard error of 0.042. Again, confidence intervals also showed that all of the rotation amounts produced regression slopes that were different from 1. As shown in Figure 8, the mean slopes in the stereomotion and combined conditions started at greater than 1 and, as the rotation amount increased, their values decreased but have never reached 1. On the other hand, in the monocular condition, the mean slopes remained below 1 for all rotation amounts.

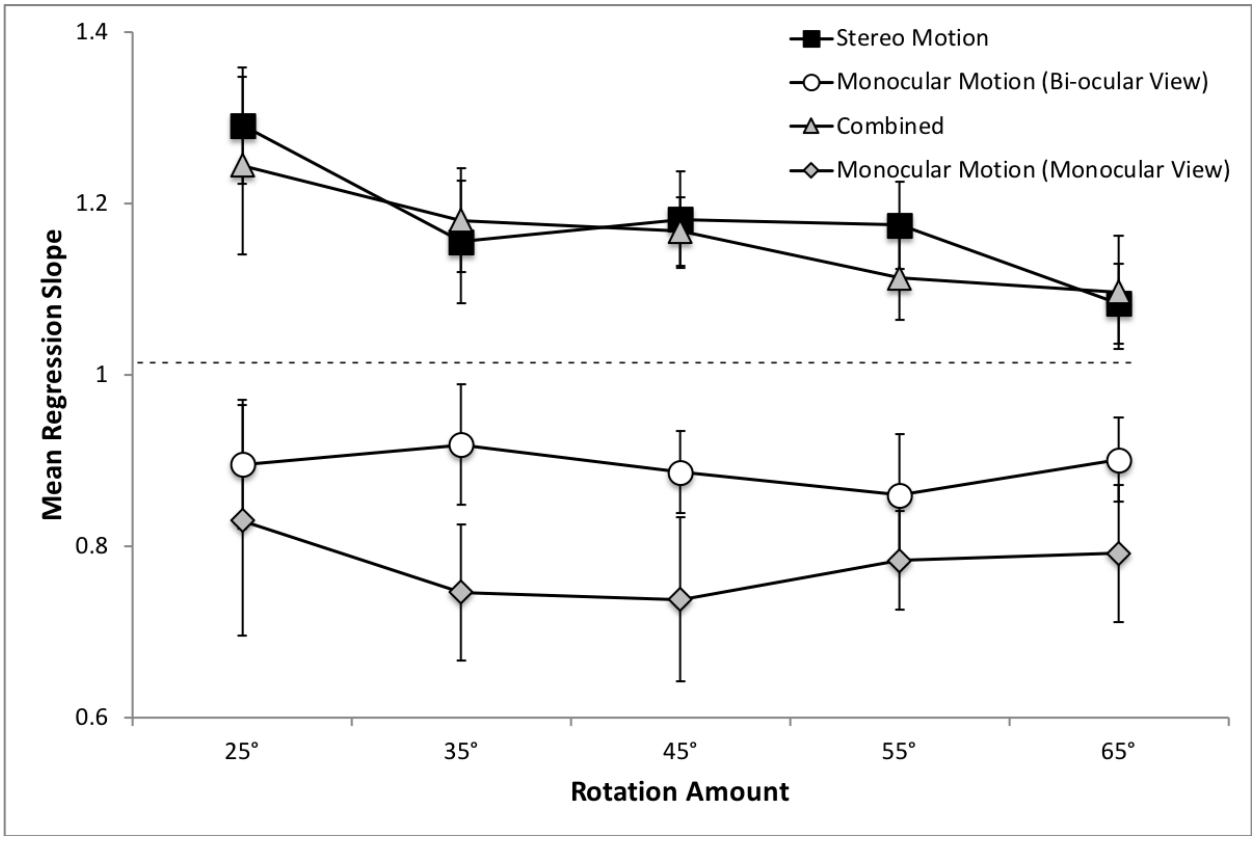


Figure 8. Mean regression slopes in Experiment 1 for the bi-ocularly viewed monocular motion condition (circles), stereo motion (squares) and combined (triangles) conditions, as well as monocularly viewed monocular motion condition (diamonds) plotted by rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

An ANOVA on regression intercepts showed that there were significant main effects of both visual information ($F(2,22) = 38.18$, $p < 0.001$, $\eta_p^2 = 0.78$), and rotation amount ($F(4,44) = 7.84$, $p < 0.001$, $\eta_p^2 = 0.42$), as well as a significant interaction effect between the two ($F(8,88) = 3.03$, $p < 0.01$, $\eta_p^2 = 0.22$). A post-hoc LSD comparison showed that intercepts for the monocular motion condition were significantly greater than those for the stereomotion condition ($p <$

0.001), and the combined condition ($p < 0.001$), whereas there was no significant difference between the two stereo conditions ($p > 0.8$). To explore the effects of rotation amount on the different visual information conditions, I used a post hoc ANOVA on the stereomotion and combined conditions. It yielded a significant main effect only of rotation amount ($F(4,44) = 10.26$, $p < 0.001$, $\eta_p^2 = 0.48$). LSD comparison showed that the 25° rotation amount exhibited significantly smaller mean intercepts than all other rotation amounts. As shown in Figure 9, for both the stereomotion and combined conditions, as the rotation amounts increased, regression intercepts also gradually increased and only reached 0 at 65° for the combined condition. An ANOVA on the intercepts in the monocular motion condition yielded no significant effects of rotation amount. However, based on their respective 95% confidence intervals, the mean intercepts for all rotation amounts were above 0.

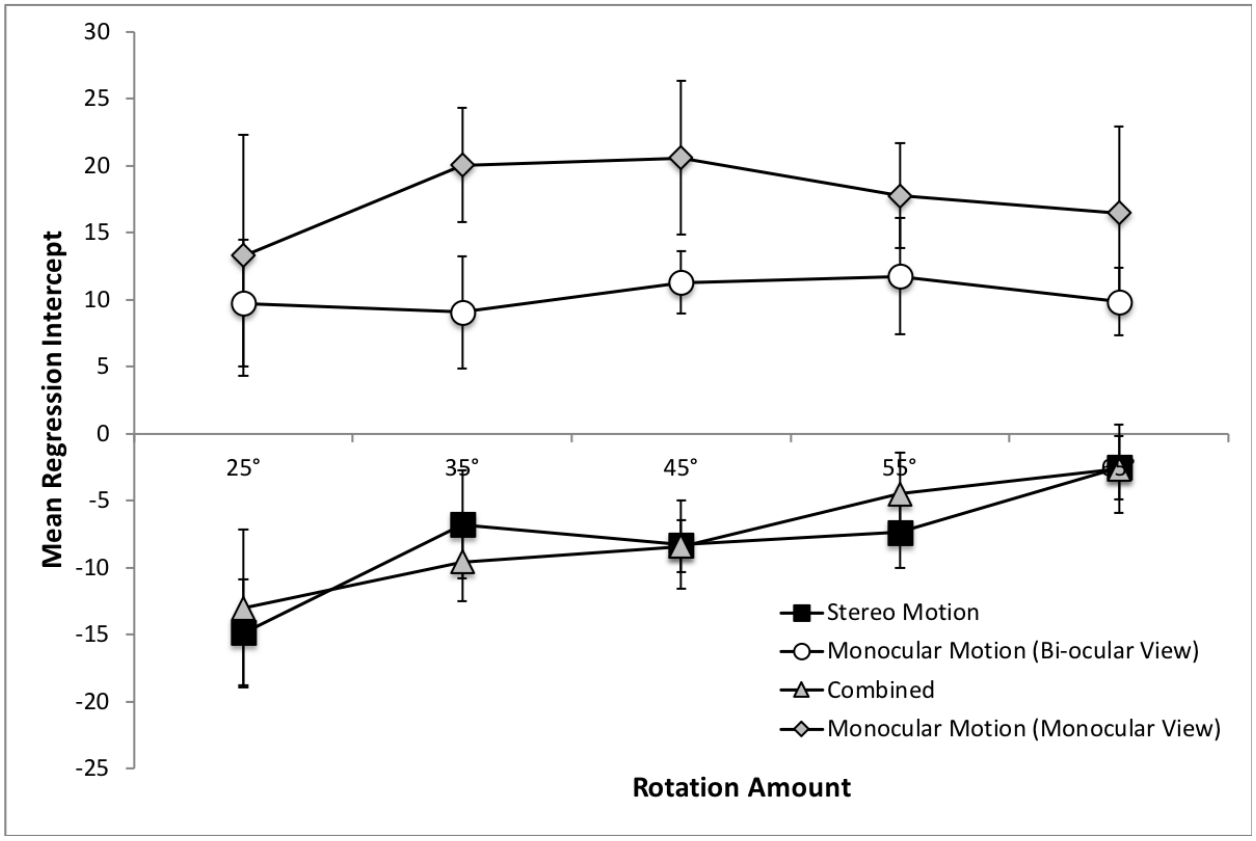


Figure 9. Mean regression intercepts in Experiment 1 for the bi-ocularly viewed monocular motion condition (circles), stereo motion (squares) and combined (triangles) conditions, as well as monocularly viewed monocular motion condition (diamonds) plotted by rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Next, an ANOVA on the r^2 yielded a significant main effect of visual information condition ($F(2,22) = 28.46$, $p < 0.001$, $\eta_p^2 = 0.72$), and of rotation amount ($F(4,44) = 5.15$, $p < 0.01$, $\eta_p^2 = 0.32$). However, there was no significant interaction effect between the two factors ($F(8,88) = 1.48$, $p > 0.1$, $\eta_p^2 = 0.12$). LSD showed that the monocular condition exhibited significantly smaller mean r^2 than both the stereomotion ($p < 0.001$) and combined conditions ($p < 0.001$), and that there was no difference between the means in the stereomotion and combined conditions. Similarly, a separate ANOVA performed only on the stereomotion and combined conditions yielded no significant main effects or interactions. Moreover, an ANOVA on the monocular condition did yield a significant main effect of rotation amount ($F(4,44) = 2.91$, $p < 0.05$, $\eta_p^2 = 0.21$). Post-hoc LSD comparison showed that the 25° rotation amount yielded a significantly smaller r^2 than did the 35° and 65° rotation conditions. As Figure 10 shows, r^2 for both the stereomotion and combined conditions remained relatively high and stable at around 0.85, while the r^2 for the monocular motion condition started off relatively low and increased as the rotation amount increased but never reached the level of the stereomotion and combined conditions.

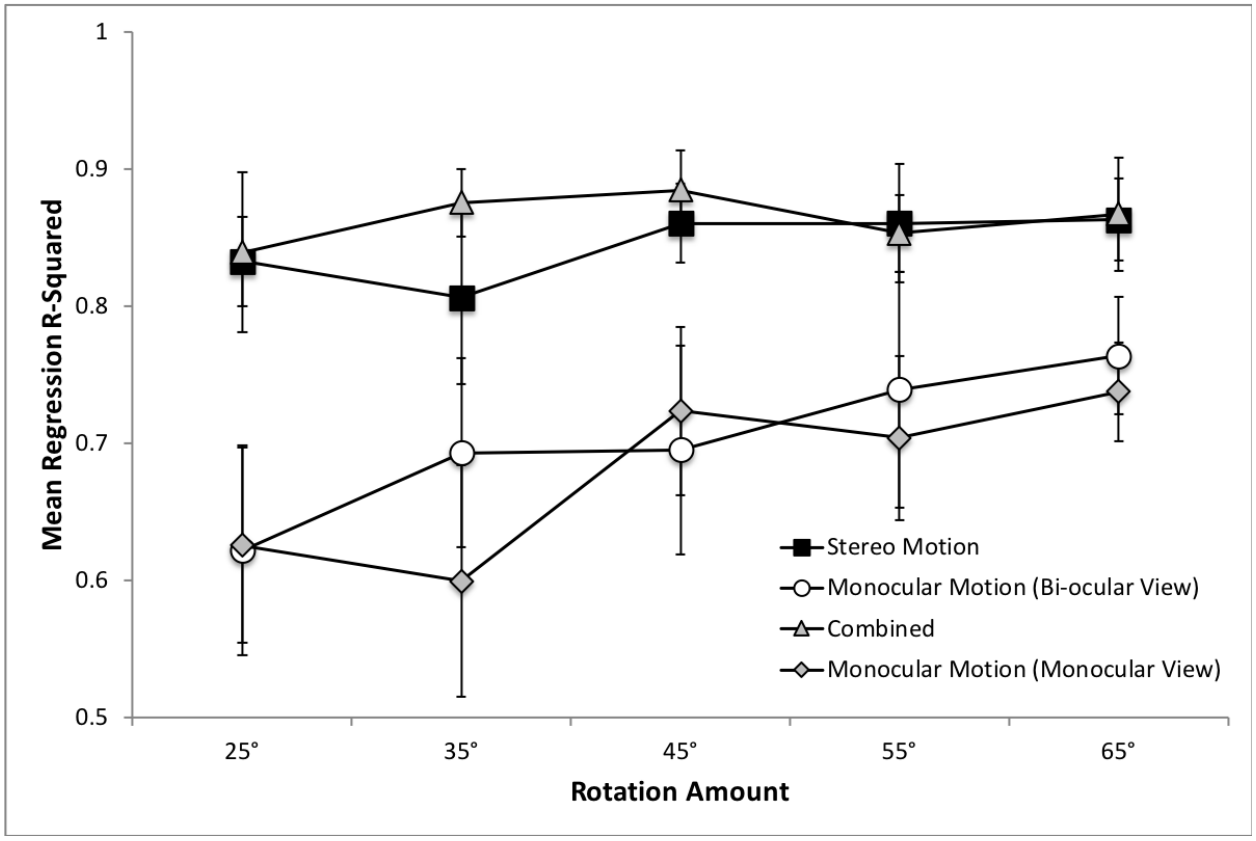


Figure 10. Mean regression R^2 in Experiment 1 for the bi-ocularly viewed monocular motion condition (circles), stereo motion (squares) and combined (triangles) conditions, as well as monocularly viewed monocular motion condition (diamonds) plotted by rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

A potential confound in Experiment 1 is that participants viewed the monocular SFM displays bi-ocularly, i.e. using two eyes. Although information regarding the slant was specified monocularly, the addition of static stereoscopic information specified the upright display screen, which might affect the judgments of slant. Thus, a control experiment was conducted using the same stimuli with 10 additional naïve participants to perform in the monocular SFM condition monocularly with their dominant eye only. None had participated in the original experiment.

Mean regression slopes, intercepts, and r^2 are plotted in Figure 8 to Figure 10. Data analysis revealed that there was no significant difference in regression slopes ($p > 0.05$), intercepts ($p > 0.1$), or r^2 ($p > 0.5$) between the two monocular information conditions. Based on confidence intervals, regression slopes were different from 1 and regression intercepts were different from 0 for all rotation amounts. Thus, when viewed monocularly, performance in the monocular motion condition was as bad or worse than when the displays were viewed bi-ocularly.

In summary, performance in the monocular SFM condition was as expected. The lack of non-coplanar points in the rigid motion of the planar surface should yield failure of structure-from-motion. Indeed, performance was inaccurate, and accuracy was unaffected by increases in the amount of perspective change. In contrast, the accuracy of judgments made using stereomotion information improved with increases in the amount of perspective change. However, the availability of stereomotion information failed to generate slopes of 1 at 45° of rotation. Instead, for the two stereomotion conditions, regression slopes only gradually decreased towards 1 but they failed to become 1 even at 65° of rotation. Such gradual change was also reflected in the change of intercepts. In the Lind et al. study, performance rather suddenly became accurate and more precise with 45° of rotation and then remained so with greater amounts of rotation.

Why should perceived slant be different in this way from perceived shape? Certainly, in the monocular SFM condition, the failure to reproduce the results of previous studies could be attributed to the lack of non-coplanar points in the planar surfaces judged in Experiment 1. Although the stereomotion conditions yielded changes in performance in response to increased perspective change, the changes did not replicate those found in previous studies. In particular, judgments failed to become accurate. The displays of strictly planar surfaces may not have

yielded perception of 3D relief structure even with stereomotion. The 3D shapes judged in previous investigations all entailed sets of rigid non-coplanar points whereas the surfaces judged in Experiment 1 did not. Therefore, in Experiment 2, I used the same paradigm to test slant perception of rectangular surfaces containing non-coplanar points.

Experiment 2

In Experiment 2, I examined the possibility that poor performance with the strictly planar surfaces was due to the lack of 3D surface structure in the display. The 2D planar surfaces might have failed to provide the viewer with the relief structure required by the bootstrap process. To test this, I added nine small 3D cuboids to the top of the rectangle surfaces. To achieve this, I cut out sections of the original surface and raised them slightly above the surface without altering the surface orientations, that is, they remained parallel to the original surface. The goal was to minimize perturbation of slant that would be confounded with the addition of non-coplanar points.

In addition, although performance in the original stereomotion and combined conditions in Experiment 1 was almost identical, the pure stereomotion condition displays failed to contain trackable texture elements, which could be understood to be either 3D or 2D (that is, optical). Specifically, to eliminate monocular information, random dots are re-randomized and re-generated in every consecutive frame (Julesz, 1971). However, stimuli in this experiment provided 3D features that could be tracked, i.e. vertices, arranged to form right angles, such as the cuboid bumps arranged in a rectangular grid. Because the original solution to the bootstrap process requires right angles, it was possible that observers could directly utilize the intrinsic 3D relief structure of the object to perform the bootstrap process despite of a lack of trackable texture elements. Therefore, the stereomotion condition in this experiment could also be used test whether trackable 3D structure, instead of trackable 2D points, could enable the bootstrap process to recover metric slant with large continuous perspective change. All other aspects of the experimental design in Experiment 2 were the same as Experiment 1. With non-coplanar points, even the monocular SFM condition was predicted to yield perception of relief 3D structure and

thus, with sufficient perspective change, perception of accurate geographical slant. Results showed improved performance in all three visual information conditions that now replicated results from the previous 3D aspect ratio judgment study.

Methods

Participants

Thirty-six adults, age between 20 and 30 (thirteen males and twenty-three females) participated in this experiment, all of who provided their informed consent prior to participating in the experiment with the approval from Indiana University's IRB. They were paid $10/hour. All participants had normal or corrected-to-normal eyesight, and also passed the same stereo fly test. There were twelve participants in each information conditions, with five males and seven females in the monocular condition, two males and ten females in the stereomotion condition, and six males and six females in the combined condition.

Stimuli and Apparatus

Experimental apparatus and display manipulations were all the same as in the previous experiment, with the exception that I added nine identical rigid cuboids on the top of each planar surface (see Figure 11 for a schematic demonstration). The top surfaces of the cuboids were parallel to the main surface and thus avoided adding noise to the display. Each cuboid was 1 cm in length and width, and 0.55 cm in height above the main surface. Depending on the slant angle, the projections of the top surface of each cuboid had a size ranging from 0.26° (23° slant) to 0.64° (73° slant). The cuboids were placed in a three-by-three grid on the top of the slanted surface and were separated by 3 cm from one another. In this way, the cuboids spanned the entire

surface. The height of the cuboids was selected so that when the surface rotated, the cuboids did not occlude one another. As with Experiment 1, the stimuli were generated using random dot anaglyphs.

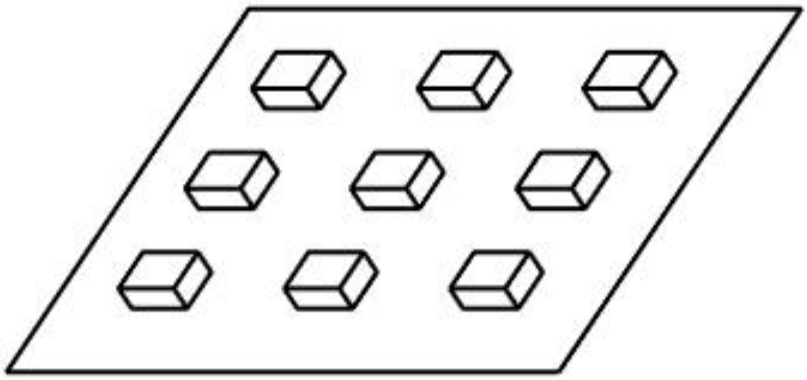


Figure 11. A schematic illustration of the non-coplanar object used in Experiment 1. Note that the actual display employs random dots.

Procedure

Unlike Experiment 1, this experiment entailed a between-subject design to provide a stronger test of each of the three information conditions. In the monocular condition, participants viewed the display monocularly with their dominant eye. Two participants' data in the combined condition were discarded due to extremely poor performance as a result of failure to correctly understand the instructions and/or lack of attentiveness during the experiment.

Results and Discussion

Multiple regression showed that surface height accounted for a very small proportion of variance for all three visual information conditions (CDOT condition, 3.5% with Cohen's f^2 = 0.11, monocular SFM condition, 4.5% with Cohen's f^2 = 0.15, and combined condition, 2.6%

with Cohen's $f^2 = 0.09$). Thus, I will not include surface height as a separate variable in the subsequent analysis.

Next, linear regressions were performed, regressing judged slant onto actual slant. Figure 12 shows the mean linear regression slopes for planar surfaces with cuboids as a function of different rotation amounts for the different visual information conditions. As the figure shows, the overall regression slopes decreased with increasing rotation in all three conditions. Changes in all conditions shifted the overall mean slopes towards 1. Mauchly's test of sphericity was not significant. ANOVA showed that there was a significant main effect of rotation amount ($F(4,124) = 10.56$, $p < 0.001$, $\eta_p^2 = 0.25$), but not a significant main effect of visual information ($p > 0.8$) or a significant interaction effect between the two ($p > 0.05$). Post-hoc comparison with LSD correction showed that regression slope at 25° rotation was significantly greater than all four other rotation amounts ($p < 0.05$) and 55° rotation had regression slopes that were significantly smaller than 35° and 45° rotation ($p < 0.05$). In fact, 95% confidence interval showed that, except for at 25° rotation, regression slopes in all other rotation conditions all included 1. As Figure 12 shows, with the addition of cuboids, performance was poor at small rotation amount but became veridical when given larger amount of rotations and in this case, at 35° of continuous perspective change.

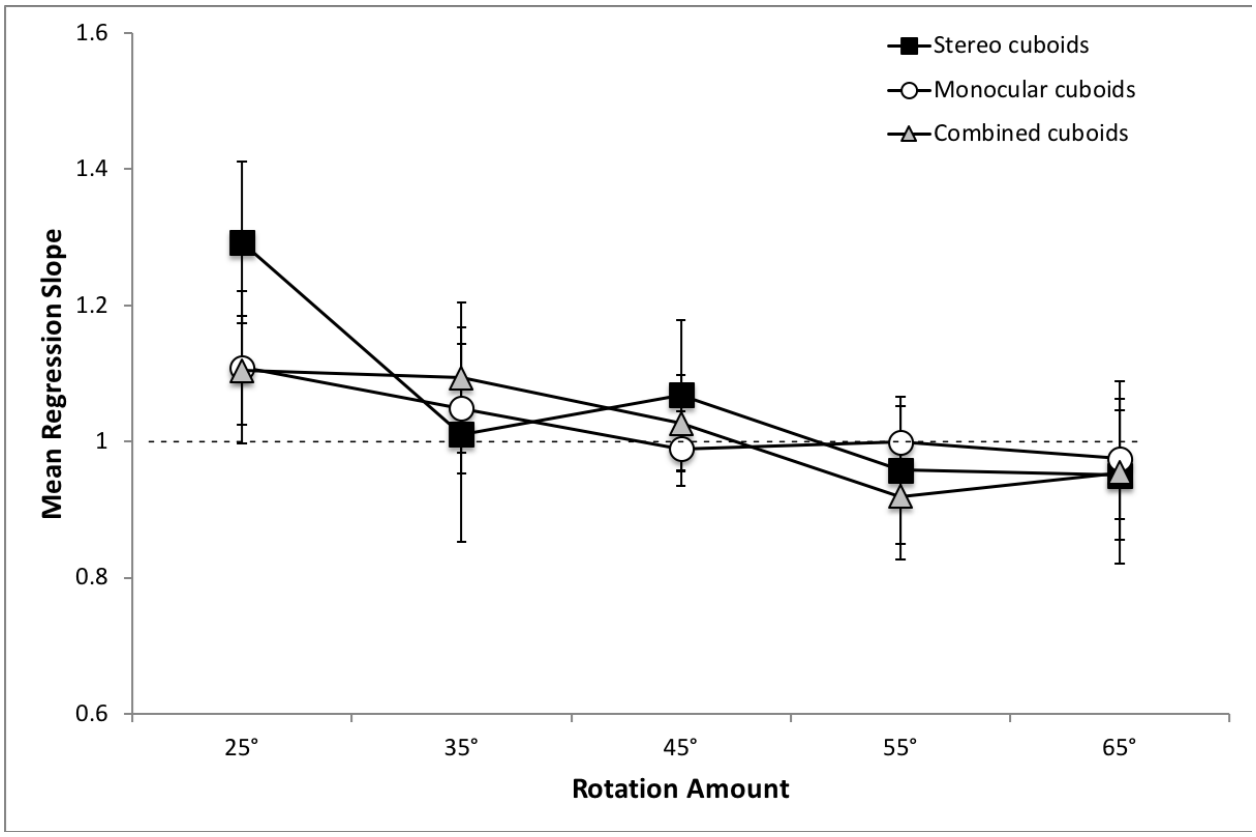


Figure 12. Mean regression slopes in Experiment *2* for the monocular motion condition (circles), stereomotion (squares) and combined (triangles) conditions plotted by rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Analyses on regression intercepts revealed a similar trend as the slopes (Figure 13). The sphericity assumption was violated ($\chi^2(9) = 17.53$, $p < 0.05$). With Greenhouse-Geisser correction, there was only a significant main effect of rotation ($F(3.02,93.61) = 9.63$, $p < 0.001$, $\eta_p^2 = 0.24$), but not a significant main effect of visual information ($p > 0.5$) or a significant interaction effect between the two ($p > 0.1$). Post-hoc LSD correction again showed that regression intercepts in 25° condition were significantly smaller than those in the other four conditions ($p < 0.05$). Intercepts in 55° condition were also significantly greater than those in 35°

rotation condition. Examining 95% confidence intervals showed that for the monocular condition, the confidence intervals included 0 for all rotation amounts except at 45° and 55° of rotation. For stereomotion condition, confidence intervals only failed to include 0 at 25° rotation. The combined condition was similar to the monocular condition, where the confidence intervals included 0 at 35°, 45°, and 65°. As shown in Figure 13, overall, increasing the amount of rotation also helped regression intercepts to be indistinguishable from 0.

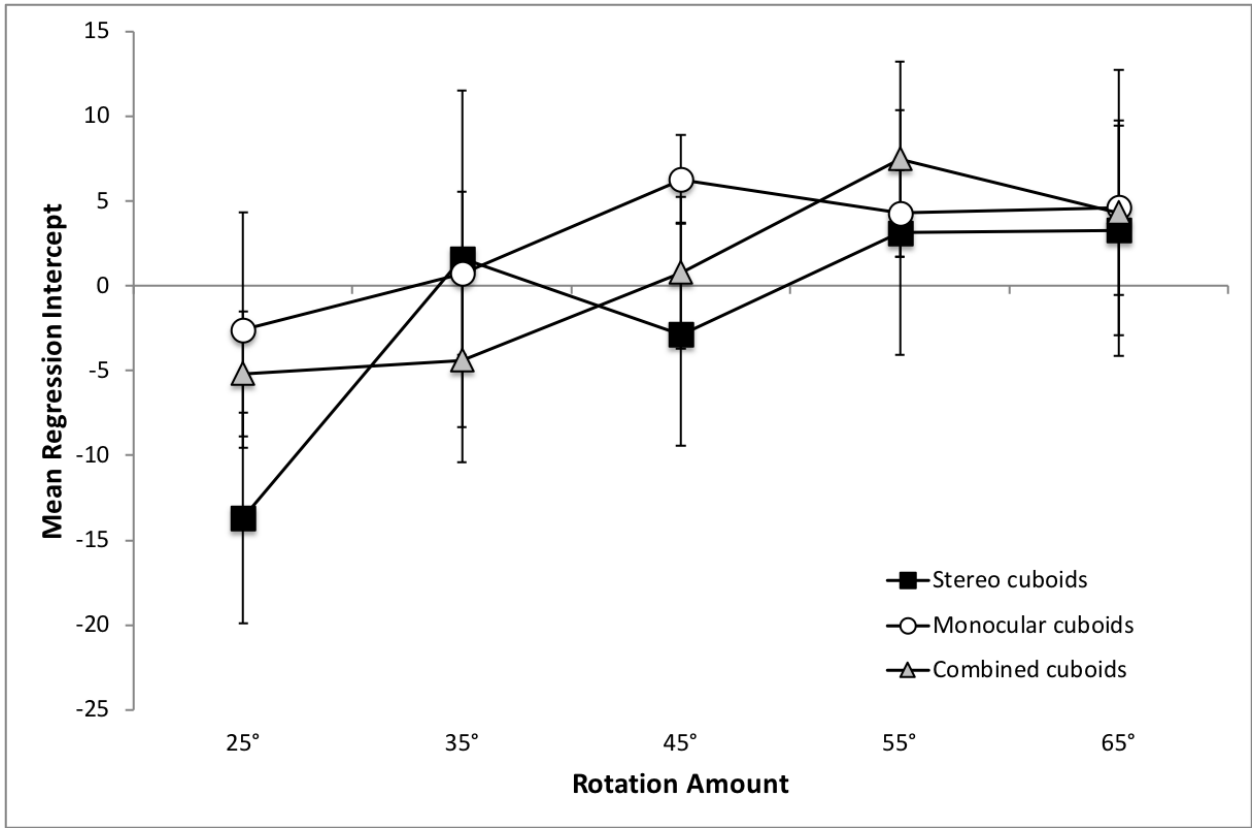


Figure 13. Mean regression intercepts in Experiment 2 for the monocular motion condition (circles), stereomotion (squares) and combined (triangles) conditions plotted by rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Finally, an ANOVA on regression r^2 showed no significant main effects of rotation amount or of visual information, or a significant interaction effect between the two. As shown in Figure 14, all three visual information conditions produced relatively stable r^2 across rotation amounts.

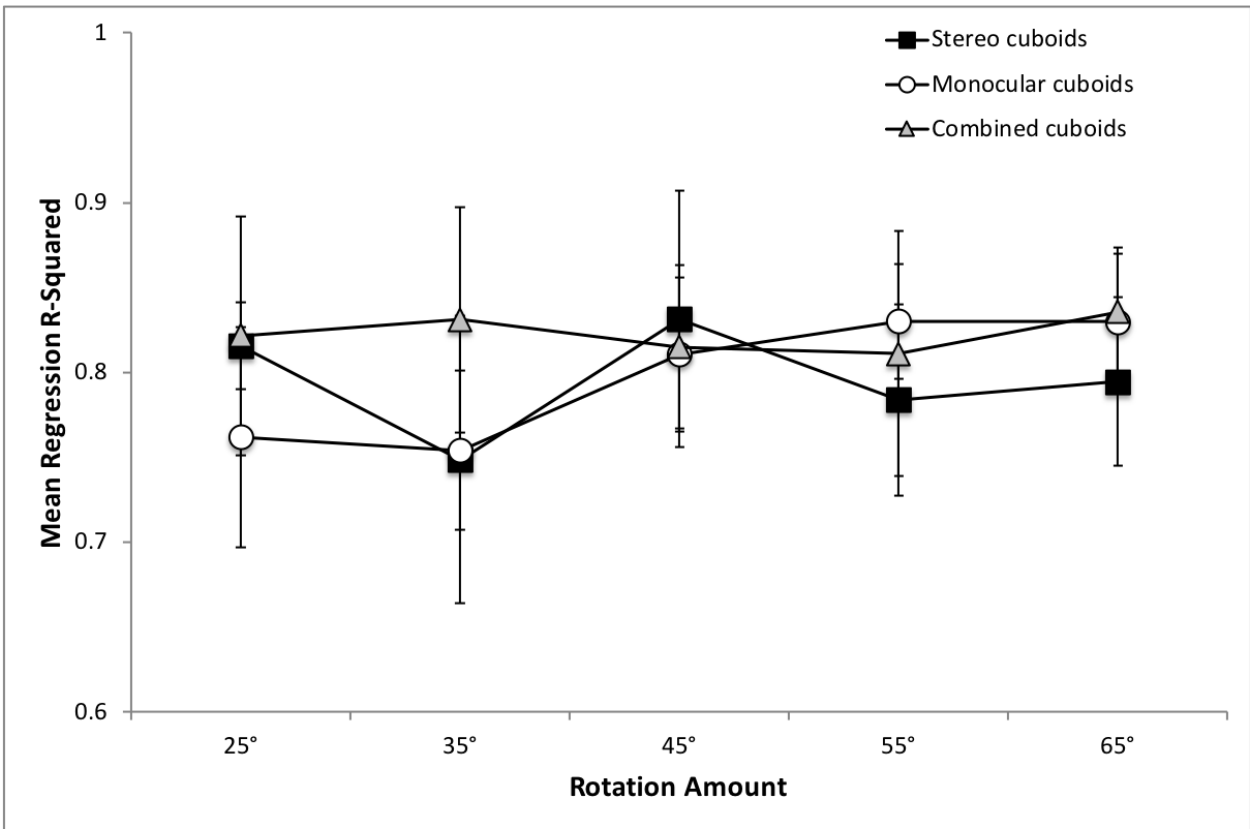


Figure 14. Mean regression r^2 in Experiment *2* for the monocular motion condition (circles), stereomotion (squares) and combined (triangles) conditions plotted by rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Results using surfaces with non-coplanar points in monocular and combined conditions in comparison with results with planar surfaces revealed that introducing additional 3D structure to planar surfaces enhanced the accuracy of slant judgments. The findings provided support for

the bootstrap process, as they replicated the pattern of results found in Lind et al. (2014) with 3D shapes. In particular, in all three visual information conditions, regression slopes were greater than 1 when the rotation amount was less than 35°, but slopes became indistinguishable from 1 and intercepts became indistinguishable from 0 once the rotation reached 35° or beyond. However, a caveat here is that we found regression intercepts in the monocular condition to be unstable as the amount of rotation increased. Although regression slopes remained at 1, intercepts overshot and went over 1 at 45° and 55° of rotation, which suggested a positive bias in these rotation conditions. Based on the adopted reference frame, such bias indicated a tendency to judge slants to be more frontoparallel, as past literature has suggested (e.g. Durgin, Li, & Hajnal, 2010; Saunders & Chen, 2015). This tendency was the most pronounced in the monocular SFM condition while absent in conditions that contained stereomotion information. This indicated that stereomotion could potentially aid the elimination of such bias, which was also consistent with previous research on this topic (e.g. Todd, Christensen, & Guckes, 2010). Nonetheless, the bias exhibited in the form of high regression slope was relatively small compared to those identified in previous studies (around 5° in the current experiment while around 15° in Durgin, Li, & Hajnal (2010)). Therefore, this will not be investigated further.

A potential confound in this Experiment concerns the necessity of motion. Specifically, it is possible that the addition of 3D cuboids provided static 3D structure that enables observers to perceive metric slant, with or without motion. To address this issue, I conducted a control experiment with stimuli used in the monocular SFM condition with cuboids and 45° of continuous perspective change, as it was when judgment became veridical. The display sampled static frames from the motion sequence when the object is frontoparallel (tilt = 0°), at either end of the perspective change (i.e. -22.5° and 22.5°), as well as the midpoints between the

frontoparallel view and the end views (i.e. -11° and 11°). Each frame was showed sequentially (from -22.5° to -11° and to 0° etc.) with a 1-second motion mask inserted between each frame, and participants had to finish viewing an entire cycle of rotation before making a judgment. When making a judgment, participants could also toggle the static frames using a computer key. Six participants were recruited, viewing the displays monocularly with their dominant eye. The mean regression slope was 0.67 (SE = 0.063), intercept 26.65 (SE = 6.85), and r^2 0.69 (SE = 0.043). A one-sampled t-test showed that the regression slopes were significantly smaller than 1 ($p < 0.01$), and intercepts were significantly greater than 0 ($p < 0.01$). Comparing with its motion counterpart, one can see that even with a reasonable sample of the static frames, motion, in this case, large continuous perspective change is still required to yield metric judgments.

The above findings indicate that relatively poor performance with planar surfaces was due to the lack of 3D structure on the surfaces preventing observers from obtaining the relief structure that would then be used to derive accurate slant estimation given increased perspective change. However, with non-coplanar elements added to the planar surface, one could use the emerging relief structure to perform the bootstrap with large rotation. Moreover, regression slopes did not monotonically decrease as a function of rotation amount. Instead, they remained relatively stable once the rotation amount was greater than 35°, further confirming the bootstrap process.

A surprising finding emerged from this experiment. The pure stereomotion condition produced similar pattern of results as the combined and monocular SFM conditions. In all three conditions, performance was poor with a small amount of rotation and became veridical as the amount of rotation increased. This suggested that the bootstrap process worked in the stereomotion condition. Because the bootstrap process hypothesizes that the perceiver tracks

points on the object to bootstrap the correct relief scaling factor and there was a lack of trackable surface texture in the pure stereomotion display, such performance could be potentially due to the particular 3D structure of the stimuli, namely, all the right angles in the object, i.e. the rectangular planar surface and the cuboids as well as the rectangular grid of the cuboids, provided 3D structure that enabled the bootstrap process. Therefore, in Experiment 3, I removed the right angles in the stimuli by using a series of symmetric hexagonal surfaces with nine tetrahedrons at random locations on the surface, presented with either pure stereomotion information or the combination of stereomotion and monocular SFM. If bootstrapping metric slant did use right angles, then one should expect consistently poor performance in the stereomotion condition but improved performance at large rotations in the combined condition where trackable surface texture could be used to find and track right angles.

Experiment 3

In Experiment 3, I eliminated intrinsic right angles formed among vertices in the object structure by substituting a hexagonal surface for the rectangular one, and tetrahedrons at random locations on the surface instead of cuboids in a rectangular grid. To further explore the effects of trackable 3D features versus trackable texture elements, I tested this setup using both pure stereomotion and the combination of stereomotion and monocular SFM.

Methods

Participants

Twenty-two adults, aged between 20 and 30, participated in this experiment. There were four males and seven females in the combined condition, and four males and seven females in the stereomotion condition. All participants provided their informed consent prior to participating in the experiment with the approval from Indiana University's Institutional Review Board (IRB). They were compensated with course credits. All participants had normal or corrected-to-normal eyesight, and also passed a stereo fly test as in previous experiments.

Stimuli and Apparatus

Experimental apparatus and setup were the same as in the previous experiment. For the stimuli, I used a symmetrical hexagon instead of a rectangle (Figure 15). To control for the potential effects of differences in size, I constructed the hexagon so that its maximum width was 10 cm, and its maximum height varied among 8 cm, 10 cm, and 12 cm. Given the viewing geometry, the horizontal dimension of the projected object was 6.7°, and its vertical dimension ranged from 2.1° to 7.7°, which was equivalent to the size of the stimuli used in Experiments 1

and 2. Additionally, I introduced nine tetrahedrons at random locations on the pentagon to avoid any explicit 90° angles among them. Each tetrahedron had a height of 0.55 cm and each triangular side had a length of 1 cm. Thus, the projected dimension of the tetrahedrons was approximately 0.37°. Additionally, the locations of the tetrahedrons were selected so that they did not block one another throughout rotation for different slant angles. I presented the stimuli with either pure stereomotion and combined information generated using the same method as in Experiments 1 and 2.

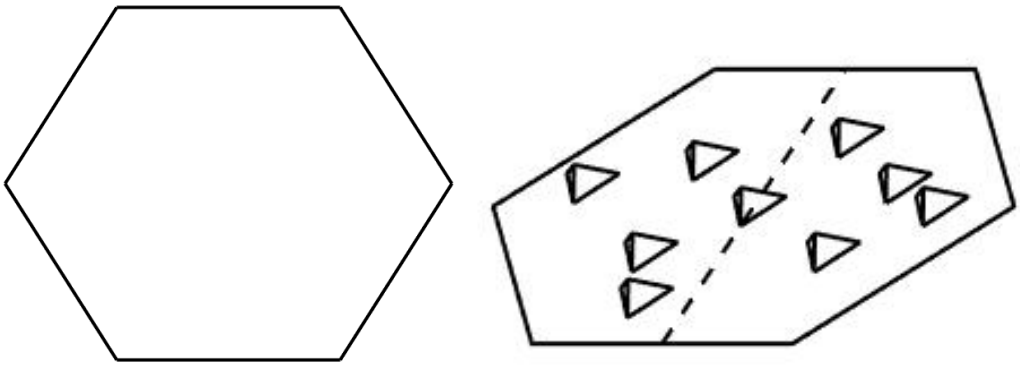


Figure 15. A schematic demonstration of the shape of the hexagon surface used in Experiment 3 (left) and layout of the actual object (right). The surface displayed here has a width of 10 cm and a height of 10 cm. The dashed line on the actual object represents the object's symmetry axis. Note that the actual displays were presented using random dots.

Procedure

The same experimental procedure was used as in previous experiments.

Results and Discussions

I first evaluated the effects of surface height using multiple regression. Surface height again only accounted for a small portion of the total variance (5.6%, with Cohen's $f^2 = 0.059$) and was not included in subsequent analysis.

For each visual information condition, I regressed judged slant onto actual slant for each participant for each rotation amount. Figure 16 to Figure 18 show the mean regression slopes, intercepts, and r^2 for the combined and stereomotion conditions. Subsequently I ran mixed-design ANOVA for the slopes, intercepts, and r^2 with rotation amount as a within-subject factor (5 levels) and visual information as a between-subject factor (2 levels). The ANOVA on regression slopes showed that there was a significant main effect of rotation amount ($F(4,80) = 5.75$, $p < 0.001$, $\eta_p^2 = 0.22$), but no significant main effect of visual information ($p > 0.9$, $\eta_p^2 = 0.00$) or a significant interaction effect ($p > 0.8$, $\eta_p^2 = 0.018$). Post-hoc analysis with LSD adjustment showed that regression slopes at 25° and 35° were not significantly different from each other but they were both significantly different from those in 45°, 55°, and 65° rotation ($p < 0.05$).

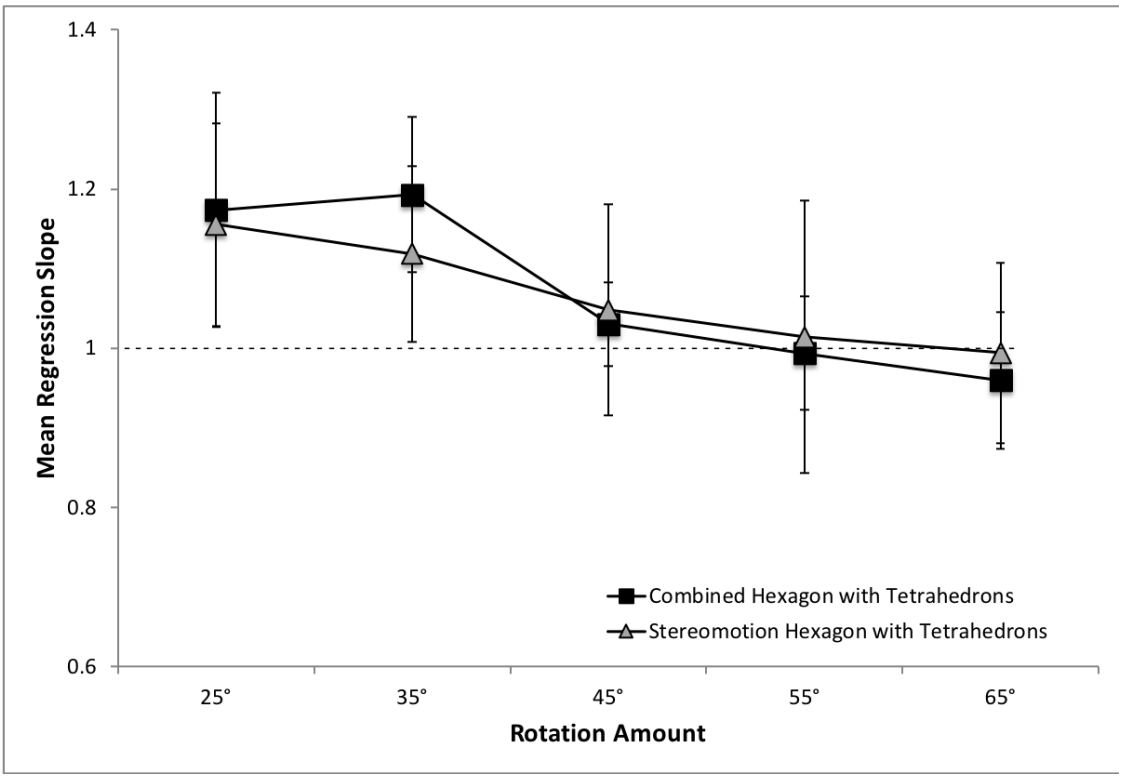


Figure 16. Mean regression slopes for the hexagon display with tetrahedrons in Experiment 2, presented using stereomotion or combined visual information. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

An ANOVA on regression intercepts showed there was a significant main effect of rotation amount ($F(4,80) = 10.06$, $p < 0.001$, $\eta_p^2 = 0.34$), but no significant main effect of visual information ($p > 0.6$, $\eta_p^2 = 0.01$) or a significant interaction effect between the two ($p > 0.3$, $\eta_p^2 = 0.06$). Post-hoc LSD correction showed that regression slopes did not differ between 25° and 35° conditions, but both rotation conditions were significantly smaller than the other three conditions ($p < 0.001$) while there was no significant difference among the latter three conditions.

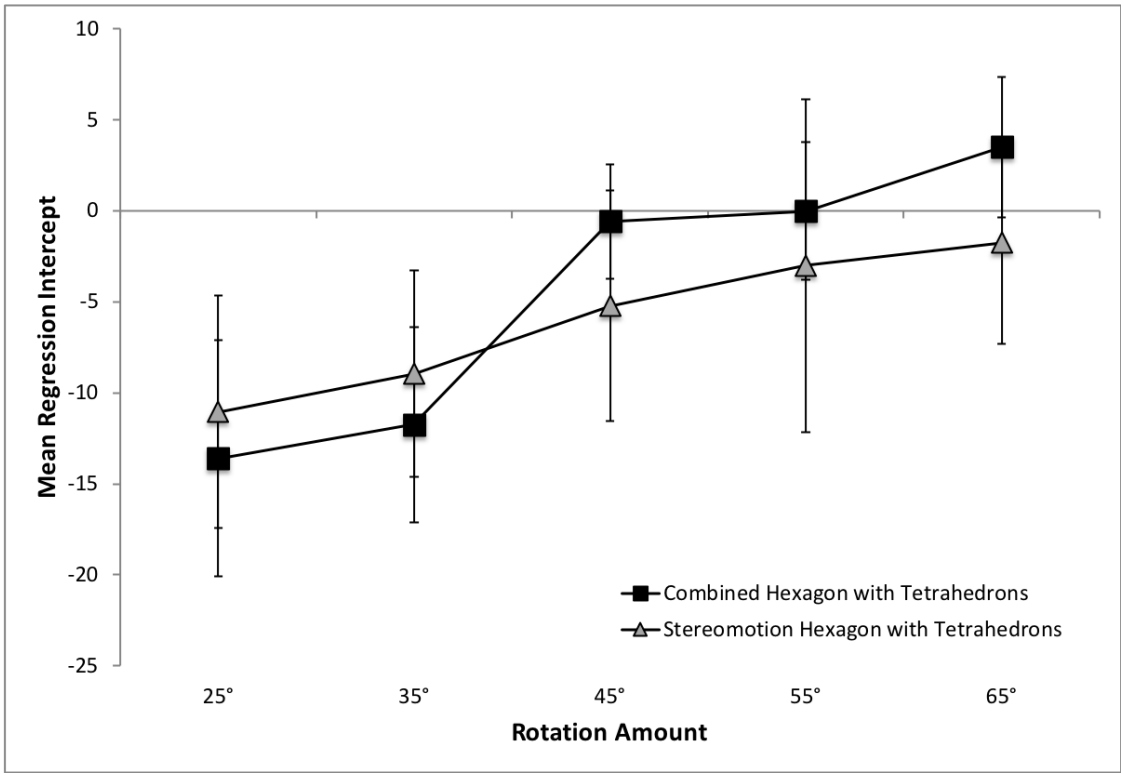


Figure 17. Mean regression intercepts for the hexagon display with tetrahedrons in Experiment 2, presented using stereomotion or combined visual information. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Next, I evaluated the quality of the performance using 95% within-subject confidence intervals around the mean for regression slopes and intercepts. As can be seen from Figure 16 and Figure 17, confidence intervals for regression slopes did not include 1 for both combined and stereomotion conditions at 25° and 35° of rotation, corresponding to intercepts not including 0 in those two conditions. This suggested that slant judgments were poor with small amounts of rotation for both stereomotion and combined conditions. However, at 45° of rotation and beyond, the confidence intervals for slopes included 1 and those for intercepts included 0. This indicated that slant judgment became veridical starting at 45° of rotation and remained relatively stable as rotation further increased.

Finally, an ANOVA on regression r^2 did not show a significant main effect for either rotation amount ($p > 0.5$, $\eta_p^2 = 0.04$) or visual information ($p > 0.1$, $\eta_p^2 = 0.10$) or a significant interaction effect ($p > 0.6$, $\eta_p^2 = 0.03$). As Figure 18 shows, there was no pronounced trend of change in r^2 as a function of rotation, where they all remained relatively stable.

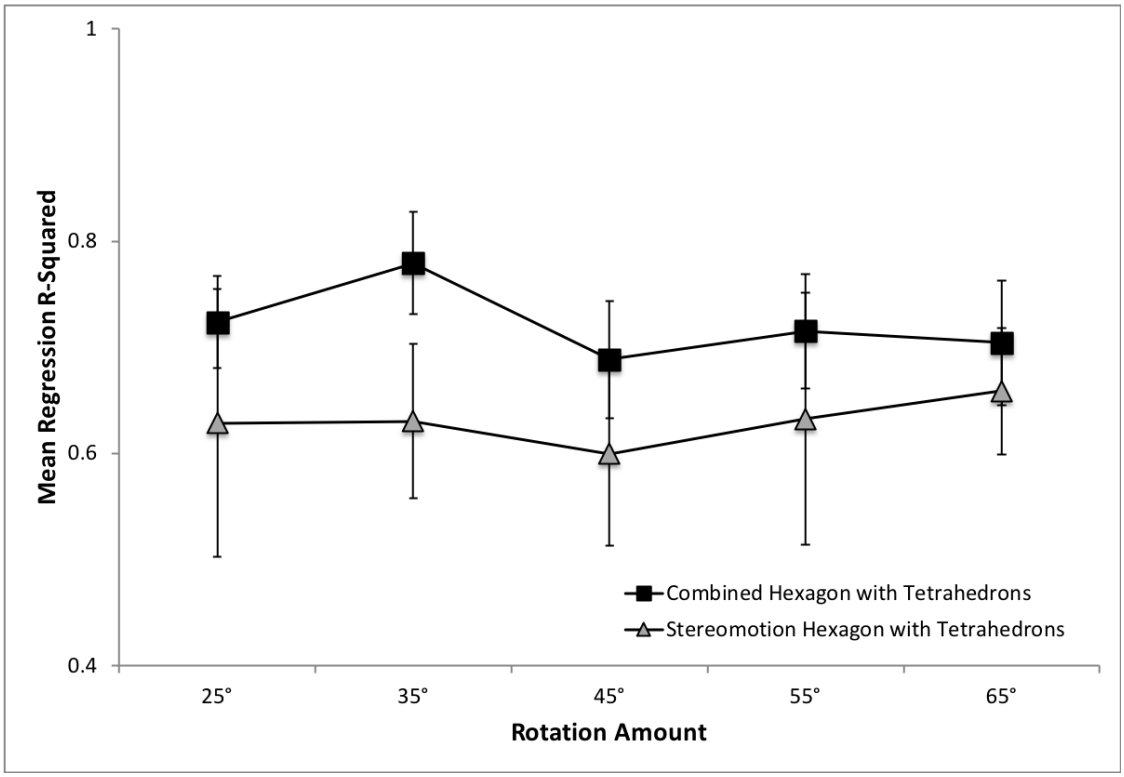


Figure 18. Mean regression intercepts for the hexagon display with tetrahedrons in Experiment 2, presented using stereomotion or combined visual information. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

In Experiment 3, I looked at the effect on slant judgments of eliminating inherent right angles in 3D structure. I constructed a symmetrical hexagon planar surface with nine tetrahedrons scattered at random locations on top of the surface to provide non-coplanar points. I

presented the stimuli using both pure stereomotion and the combination of stereomotion and SFM. With pure stereomotion, there were no trackable 3D texture elements that could potentially form right angles or any right angles inherent to the object. If participants used the right-angle solution of the bootstrap process to perceive metric slant, then one should expect poor performance in the stereomotion condition and good performance in the combined condition where there was trackable texture that would allow formation of right angles. However, there was no difference between the two visual information conditions. For both conditions, performance was poor with small amounts of rotation, but became veridical at 45° of rotation and beyond. The difference between the small and large rotation conditions showed that the bootstrap process was performed with this set of stimuli, but not necessarily with the right-angle solution. So, what could the bootstrap process be here?

Upon a closer examination of the original right-angle solution, I noted that the equidistant points that made up one leg of the right angle could potentially be used to extrapolate the relief scaling factor. By definition, each pair of equidistant points are in a frontoparallel plane and the distance between them can be accurately perceived. As the object rotates, the two points would no longer be in a frontoparallel plane and their distance would be subject to relief scaling. Following a similar rationale as presented in the original right-angle solution, observers can track the 3D distance between these points to bootstrap the correct scaling factor. This new solution to the bootstrap process will be further developed in Chapter III.

Following the generally positive results from applying the bootstrap paradigm to 3D slant perception, the next question to address is whether symmetry plays a role in bootstrapping metric slant. I noted that both types of stimuli in Experiments 2 and 3 contained mirror symmetry (rectangles and hexagons). Work by Pizlo and colleagues showed that symmetry could be a

constraint on perceiving 3D objects (Pizlo, 2010; Pizlo et al., 2010; Li et al., 2011). In addition, Saunders and Knill (2001) demonstrated that symmetry and the resulting skew symmetry after projection play a role in the perception of local slants when given 2D images of the surface. Because combining a geographical frame of reference and large continuous perspective change would render the direction of slant, i.e. geographical tilt, to be ambiguous, it is possible that observers in the experiments thus far used symmetry to establish the direction of slant. Therefore, in Experiment 4, I eliminated mirror symmetry in the stimuli and explored whether slant judgments would change accordingly.

Experiment 4

In Experiment 4, I removed mirror symmetry in the object by using an asymmetric pentagonal surface, scattered with the same tetrahedrons at random locations as in Experiment 3. I again presented the stimuli using either pure stereomotion or the combination of stereomotion and monocular SFM.

Methods

Participants

Twenty adults participated in this experiment. There were four males and six females in the combined condition, and four males and six females in the stereomotion condition. All participants provided informed consent prior to participating in the experiment with the approval from Indiana University's Institutional Review Board (IRB). They were either paid $10/hour or compensated with course credits. All participants had normal or corrected-to-normal eyesight, and also passed a stereo fly test as in previous experiments.

Stimuli and Apparatus

Experimental apparatus and setup were the same as in the previous experiment. For the stimuli, I used an asymmetrical pentagon as the base of the slanted object. The pentagon had a maximum width of 10 cm and maximum height varied among 8 cm, 10 cm, and 12 cm. Given the viewing geometry, the horizontal dimension of the projected object was 6.7°, and its vertical dimension ranged from 2.1° to 7.7°. The same tetrahedrons as in Experiment 3 were added to the surface as a means to introduce non-coplanar points. They were set at random locations so that to avoid right angles in the stimuli. Figure 19 offers a schematic illustration of the shape of the pentagon used in this experiment. The stimuli were presented either with only pure stereomotion

or with the combination of stereomotion and monocular SFM generated using the same method as in Experiment 3.

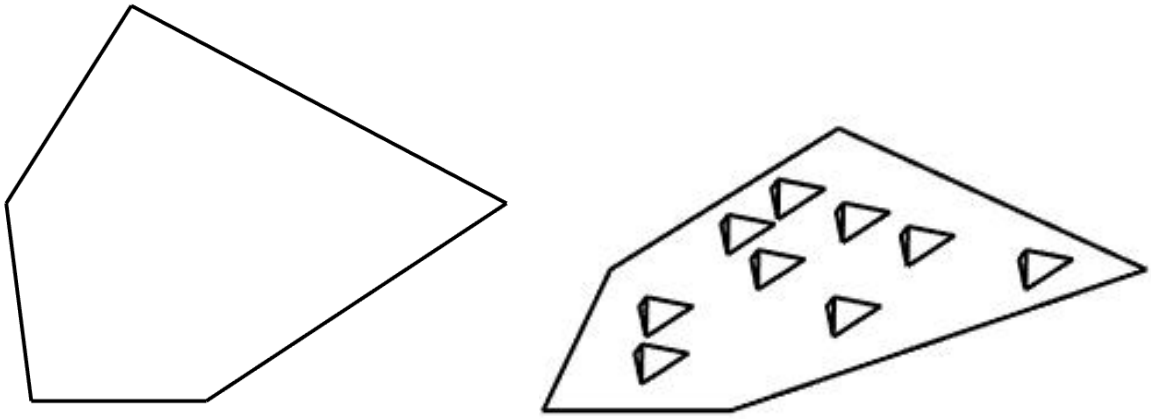


Figure 19. A schematic demonstration of the shape of the pentagon surface used in Experiment 2 (left) and layout of the actual object (right). The surface displayed here has a width of 10 cm and a height of 10 cm. Note that the actual displays were presented using random dots.

Procedure

I used the same experimental procedure as in previous experiments.

Results and Discussions

The initial multiple regression showed that the surface height accounted for a small portion of the total variance (1.4%, with Cohen's $f^2 = 0.01$), and consequently was not be included in subsequent analysis.

I performed linear regressions for each participant within each rotation amount and each visual information condition, and evaluated performance using regression slopes, intercepts, and r^2 (Figure 20 to Figure 22). An ANOVA on regression slopes showed that there was a significant

main effect of rotation amount ($F(4,72) = 10.57$, $p < 0.001$, $\eta_p^2 = 0.37$), but not of visual information ($F(1,18) = 3.07$, $p = 0.97$, $\eta_p^2 = 0.15$) or of the interaction ($p > 0.9$, $\eta_p^2 = 0.01$). Pairwise comparison showed that slopes in the 25° condition were significantly greater than those in other conditions except for at 35° ($p < 0.05$), while 65° rotation condition were significantly smaller than those in other conditions ($p < 0.05$).

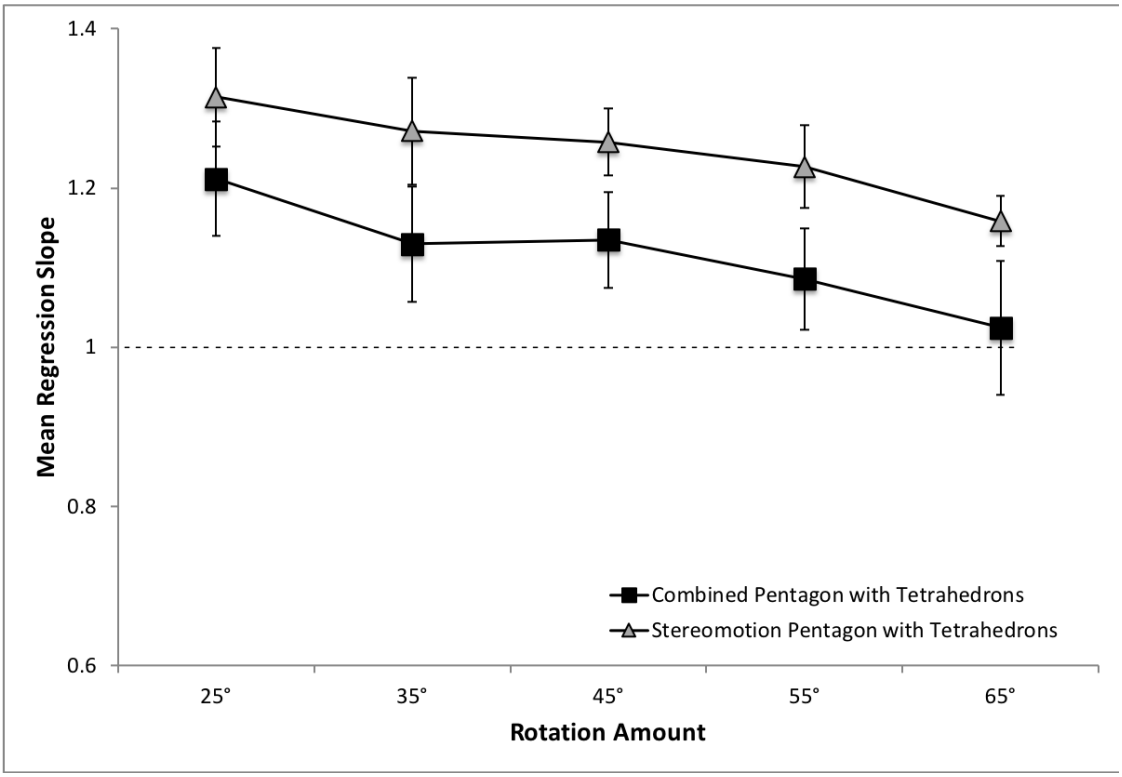


Figure 20. Mean regression slopes in Experiment 3 for the slanted pentagonal surface with nine tetrahedrons at random locations on the surface displayed with either stereomotion or the combination of monocular SFM and stereomotion information, plotted as a function of rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

An ANOVA on regression intercepts showed that there was a significant main effect of rotation amount ($F(4,72) = 13.26$, $p < 0.001$, $\eta_p^2 = 0.42$). However, there was no main effect of visual information ($p > 0.4$, $\eta_p^2 = 0.04$) or a significant interaction ($p > 0.9$, $\eta_p^2 = 0.01$). Pairwise comparisons showed that intercepts at 25° of rotation were significantly smaller than those in the other conditions ($p < 0.05$) and intercepts at 65° rotation were significantly greater than those in other conditions ($p < 0.05$).

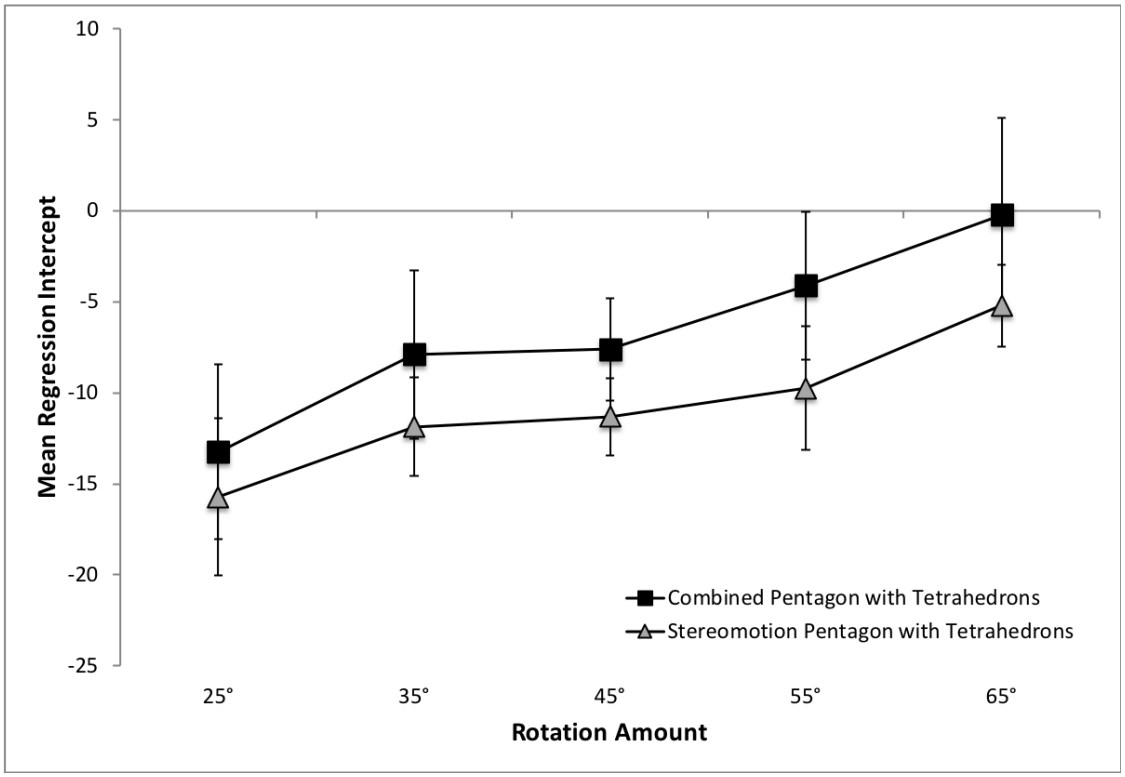


Figure 21. Mean regression intercepts in Experiment 3 for the slanted pentagonal surface with nine tetrahedrons at random locations on the surface displayed with either stereomotion or the combination of monocular SFM and stereomotion information, plotted as a function of rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

To evaluate the accuracy of performance, the 95% within-subject confidence interval around the mean was examined for regression slopes and intercepts. As shown in Figure 20 and Figure 21, regression slopes did not reach 1 and regression intercepts did not reach 0 with as much as 55° of rotation for both stereomotion and combined conditions. At 65° rotation, regression slopes reached 1 and intercepts reached 0 only for the combined condition, but not for the stereomotion condition. This was in stark contrast with the pattern of results in the previous experiments, where performance became veridical (i.e. regression slope equaled 1 and intercept equaled 0) at either 35° or 45° rotation and remained so with subsequent increase in rotation. In other words, slant judgment with the pentagonal objects remained relatively poor throughout rotation. However, there were some noticeable changes in performance as a function of increasing rotation amount, as seen from the ANOVA results.

Finally, the ANOVA on r^2 yielded no significant main effect of rotation amount ($p > 0.05$, $\eta_p^2 = 0.12$), of visual information ($p > 0.7$, $\eta_p^2 = 0.005$), or an interaction ($p > 0.9$, $\eta_p^2 = 0.01$). As shown in Figure 22, regression r^2 remained relatively stable for both visual information conditions.

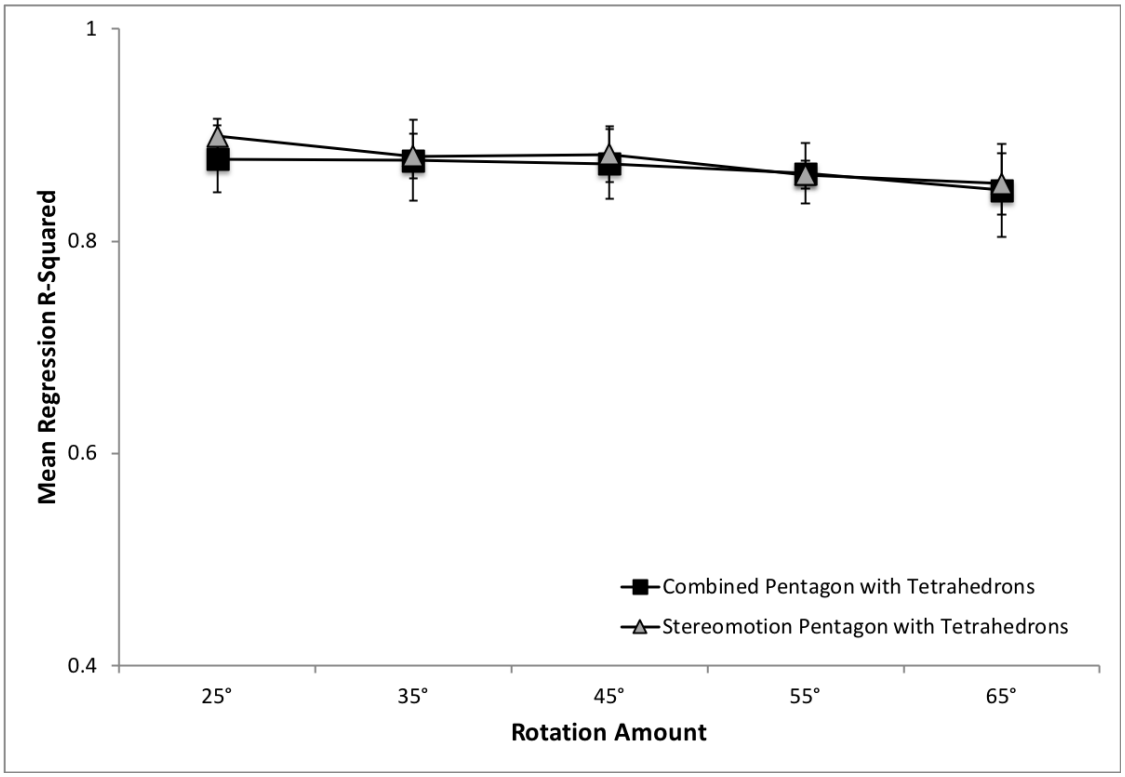


Figure 22. Mean regression r^2 in Experiment 3 for the slanted pentagonal surface with nine tetrahedrons at random locations on the surface displayed with either stereomotion or the combination of monocular SFM and stereomotion information, plotted as a function of rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Results from Experiment 4 did not replicate the pattern that was expected from the bootstrap process. This failed for both the stereomotion and the combined conditions. Such a failure could be due to two reasons. First, it could be because the bootstrap process failed to extrapolate the correct scaling factor. With everything being equal, the only difference between Experiments 3 and 4 was the presence of mirror symmetry in the object. In other words, this interpretation would suggest a necessary role of mirror symmetry in the bootstrap process.

However, Lind et al. (2014) used asymmetric polyhedrons to test aspect ratio judgments, where the top cross section of the polyhedrons shared almost identical shapes as the pentagons used in the current experiment. Judgment in their study still became veridical when there was sufficiently large continuous perspective change. This indicated that the first interpretation was unlikely. The alternative explanation is that the direction of slant was based on the symmetry axis and the lack of symmetry in the current experiment led to ambiguity in the direction of slant. Although observers could potentially recover the correct scaling factor with large continuous perspective change, this factor was paired with some random tilt that was different from the actual tilt. To further explore the role of a symmetry axis, in Experiment 5, I perturbed the symmetry axis of the hexagonal object used in Experiment 3.

Experiment 5

In Experiment 5, I used the hexagonal objects in Experiment 3 and rotated them around their surface normal by 15°. This is equivalent to changing the roll of the object (or spin as described by Saunders and Knill (2001)). If the symmetry axis was used to identify the direction of slant, regression slopes and intercepts derived by regressing judged slants onto actual slant should first exhibit a similar pattern as one would expect from the bootstrap process, i.e. regression slopes would decrease and intercepts would increase as the amount of rotation increased, and remained relatively stable once rotation reached and went beyond 45°. However, because the judged slant was based on the direction of the symmetry axis, the actual values of slopes and intercepts should be off. Alternatively, if regression was based on slant specified by the symmetry axis, regression slopes should become 1 and intercepts, 0, at 45° of rotation and beyond. Because of the lack of difference between stereomotion and combined information in Experiment 3, we only presented the stimuli using combined visual information.

Methods

Participants

Ten adults participated in this experiment, with four males and six females. All participants provided informed consent prior to participation with the approval from Indiana University's Institutional Review Board (IRB). They were compensated with course credits. All participants had normal or corrected-to-normal eyesight, and also passed a stereo fly test as in previous experiments.

Stimuli and Apparatus

Experimental apparatus and setup were the same as in Experiment 2. I rotated the object around its normal by 15° (Figure 23). Consequently, the angle between the direction of slant and the symmetry axis was 15°. The object (Figure 23 right) had the same slant and tilt as that presented in Figure 15 (right).

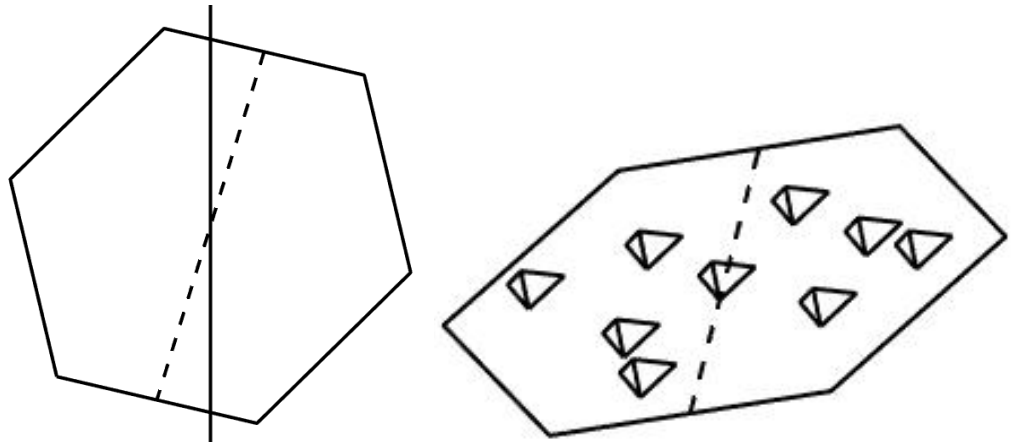


Figure 23. A schematic demonstration of the hexagon shape rotated around its normal by 15° (left) and the layout of the actual object (right). The dashed line shows the symmetry axis of the surface and the solid line shows the direction of slant.

Procedure

I used the same experimental procedure as in previous experiments.

Data Analysis

I used the same analysis protocol as in previous experiments. In addition, I also "corrected" the actual slant by computing slant in the direction of the symmetry axis. I conducted separate linear regressions using the uncorrected and corrected slant. Subsequently, I used mixed-design ANOVA's to analyze the resulting regression slopes and intercepts with five levels

of within-subject factor of rotation amount and two levels of type of slant, either corrected based on the direction of the symmetry axis or uncorrected based on the direction of surface normal.

Results and Discussion

Figure 24 shows the mean regression slopes based on uncorrected and corrected slants. Mauchly's test of sphericity showed that the sphericity assumption was violated ($\chi^2(9) = 25.20$, $p < 0.01$). With Greenhouse-Geisser correction, ANOVA on regression slopes showed a significant main effect of rotation amount ($F(2.54,45.66) = 18.36$, $p < 0.001$, $\eta_p^2 = 0.51$), but not a significant main effect of the type of slant ($p > 0.1$, $\eta_p^2 = 0.13$) or a significant interaction effect ($p > 0.9$, $\eta_p^2 = 0.002$). As shown in Figure 24, before correction, although regression slopes did exhibit similar patterns as one would expect with an increasing amount of rotation, namely that the slopes dropped as rotation increased to 45° and remained relatively steady with further increase of rotation, the actual values of the regression slopes plateaued at a level below 1 (the upper bound of 95% confidence interval was below 1 at 55° and 65° rotations). With the slant determined along the symmetry axis (that is, "corrected"), the regression slopes actually became 1 at 45° of rotation and remained steady.

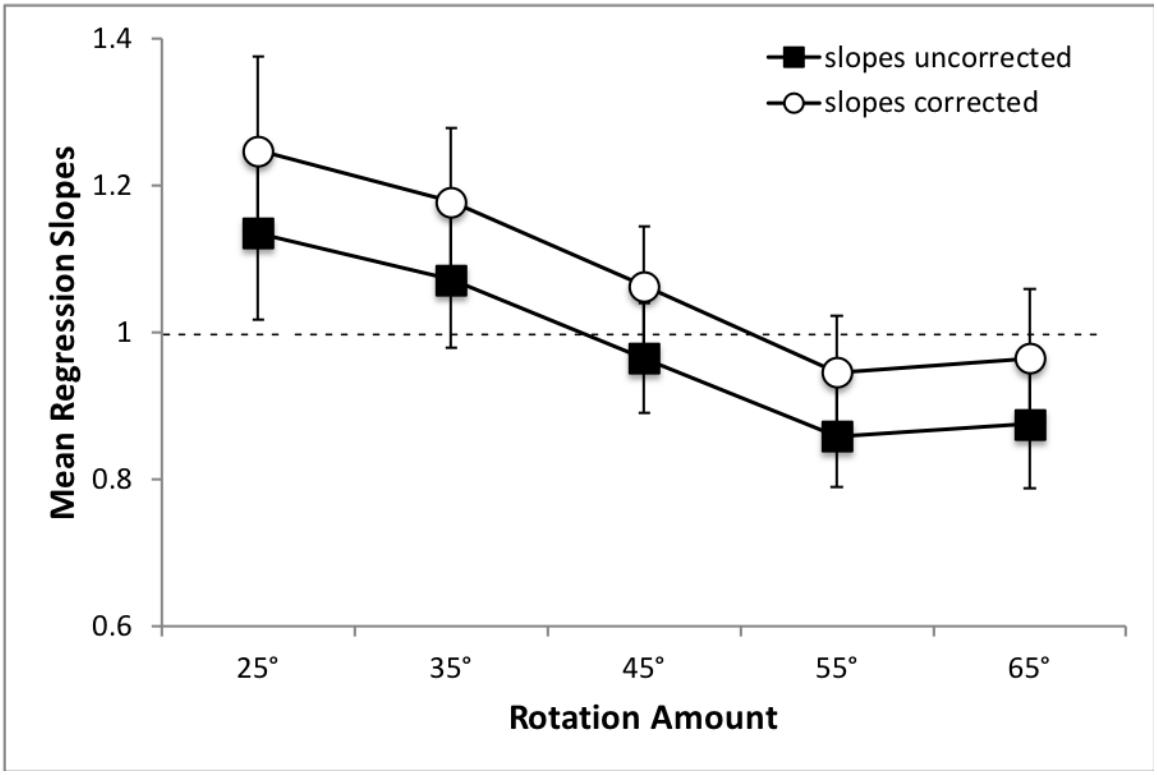


Figure 24. Mean regression slopes computed using uncorrected actual slant and corrected slant based on the direction of the symmetry axis for the hexagonal objects with a roll of 15° presented using combined information. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Regression intercepts showed a similar pattern, where an ANOVA showed that there was a significant main effect of rotation amount ($F(4,72) = 24.94$, $p < 0.001$, $\eta_p^2 = 0.58$), but not a significant main effect of type of slant ($p > 0.5$, $\eta_p^2 = 0.023$) or a significant interaction effect ($p = 1.0$, $\eta_p^2 = 0$). As shown in Figure 25, regression intercepts based on uncorrected and corrected slants were almost identical, but those from the uncorrected were slightly higher than the corrected. Comparing confidence intervals at each rotation amount, regression intercepts for both

types of slant were actually indistinguishable from 0 at 45° of rotation and remained there as rotation increased to up to 65°.

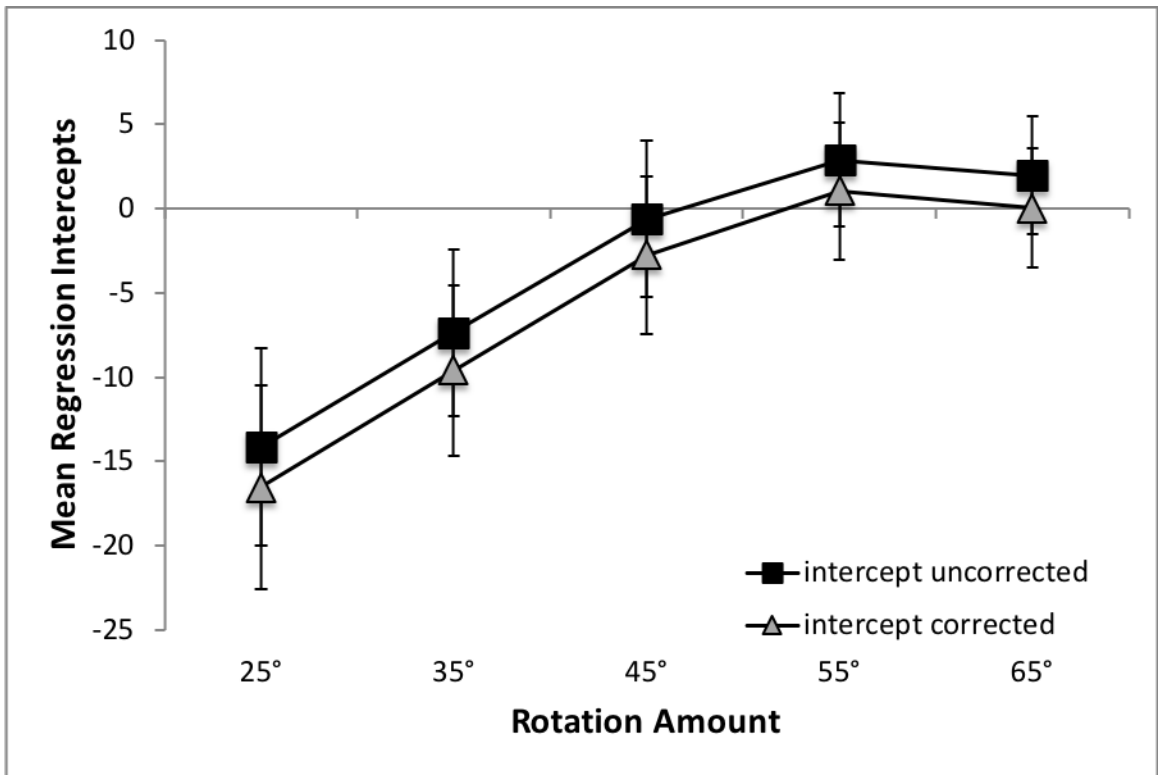


Figure 25. Mean regression intercepts computed using uncorrected actual slant and corrected slant based on the direction of the symmetry axis for the hexagonal objects with a roll of 15° presented using combined information. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

This experiment confirmed the second postulate regarding the role of symmetry, namely that symmetry axis is used to determine the direction of slant. This also implies that the lack of symmetry of the pentagonal object does not necessarily mean that people cannot obtain the

correct scaling factor with large amounts of perspective change. Instead, the observers could get the correct scaling factor with the pentagonal objects, but that factor was used to recover the wrong slant due to the ambiguity in the direction of slant. This claim will be further substantiated in Chapter III using model simulation.

Experiment 6

From Experiments 1 to 5, slant perception was tested using various kinds of planar surfaces, with or without additional 3D structure. This type of object itself is rather rare in real life situations: one does not see or interact with a floating, surface in midair with bumps on it. From a functionalist's perspective, it is essential to generalize the results based on stimuli that came from laboratory fabrication and move to more ecologically valid stimuli. Therefore, in Experiment 6, I used a series of 3D generalized cone objects that were akin to something that one might encounter in real life. These objects varied in terms of their cross-sections, all of which were used in the previous experiments, namely rectangles, pentagons, and hexagons. An additional benefit of using these objects was to further enhance of the power of the bootstrap process. Previous experiments employed small 3D structures on top of the slanted surface as a means to introduce non-coplanar points for the SFM process. With the generalized cone objects, the slanted top surface remained smooth and the entire object provides the required 3D structure to obtain the relief structure.

Methods

Participants

Thirty adults participated in this experiment. There were five males and five females in the rectangle condition, three males and seven females in pentagon condition, and two males and eight females in the hexagon condition. All participants provided informed consent prior to participation with the approval from Indiana University's Institutional Review Board (IRB). They were compensated with course credits. All participants had normal or corrected-to-normal eyesight, and also passed a stereo fly test as in previous experiments.

Stimuli and Apparatus

A series of generalized cones with different cross sections were used to integrate slant with 3D polyhedrons. These objects were generated following translational symmetry (Pizlo et al., 2010). Specifically, for instance, a generalized cone with a pentagon cross section can be generated by "sweeping" the 2D pentagon surface through space along a certain path that (following Pizlo) forms the spine of the object. This ensures that the cross section of the 3D object will be the same all along the path. For the purpose of current experiment, a polyhedral object with a flat top was first specified. Then, its top surface was rotated around its front edge by a certain angle so that the top surface is slanted at a certain angle while retaining the cross section of the object.

Both the slant of the top surface and depth-to-width aspect ratio of the cross section of the generalized cones were manipulated. Similar to previous experiments, slant was defined as the angle formed between the normal of the top surface and the y-axis, which was the axis of rotation that produces perspective change. There were again 24 different slant angles, from 27° to 73° with 2° increment between each. Variations in aspect ratio were yielded changes in the slanted surface height similar to previous experiments. There were three different aspect ratios, 0.8, 1.0, and 1.2. The maximum width of the cross sections was maintained at 10 cm. Therefore, a 0.8 aspect ratio corresponds to the depth of 8 cm and 1.2 corresponds to 12 cm.

Procedure

This experiment followed the same procedure as previous experiments with one exception. To avoid having too many trials with a complete crossing between aspect ratio and

slant angle, the 24 slants were divided into 3 groups and assigned to three different aspect ratios. Unlike previous experiments, assigning slants to each aspect ratio was done systematically instead of randomly. For instance, the aspect ratio of 0.8 corresponded to slants of 27°, 33° and so on, 1.0 to 29°, 35° and so on, and 1.2 to 31°, 37° and so on. This step was to ensure slant was sampled in a fixed interval for each aspect ratio. There was a total of five different rotation amounts, from 25° to 65° with 10° increment. Within each rotation amount, I blocked trials by aspect ratio: objects that had the same aspect ratio were presented together. Slants were randomized within each block of aspect ratio, and the order of aspect ratio was also randomized within each block of rotation amount.

Results and Discussions

Similar to treatment of surface height in previous experiments, multiple regressions were used to evaluate the effects of aspect ratio on judged slant for each type of surface individually with rotation amount, aspect ratio, and slant as predictors and judged slant as the dependent variable. For all three types of surfaces, aspect ratio contributed to a small portion of the variance (2.9% for the pentagonal cross-section with Cohen's $f^2 = 0.030$, 1.9% for the hexagonal cross-section with $f^2 = 0.019$, and aspect ratio did not satisfy the inclusion criterion for the rectangular cross-section, $p > 0.1$). Therefore, aspect ratio was not included in subsequent analyses.

Mixed-design ANOVA's were used for regression slopes, intercepts, and r^2, respectively, with rotation amount as a within-subject factor (5 levels) and shape of the object's cross-section as a between-subject factor (3 levels). ANOVA on slopes showed that the sphericity assumption was violated ($\chi^2(9) = 37.91$, $p < 0.001$). With Greenhouse-Geisser correction, there was a significant main effect of rotation amount ($F(2.12,57.16) = 18.33$, $p < 0.001$, $\eta_p^2 = 0.40$).

However, there was no significant main effect of the shape of cross section ($p > 0.2$) or a significant interaction effect between rotation and shape ($p > 0.3$). Post-hoc LSD correction showed that slopes at 25° of rotation were significantly greater than those in other rotation conditions ($p < 0.01$). Slopes at 35° rotation condition were significantly greater than those in the 55° condition ($p < 0.001$) and 65° ($p < 0.001$), but were only marginally greater than those at 45° ($p = 0.06$). Slopes in the 45° rotation condition were significantly greater than those in 55° and 65° rotations ($p < 0.01$) and there was no significant difference between 55° and 65° rotations.

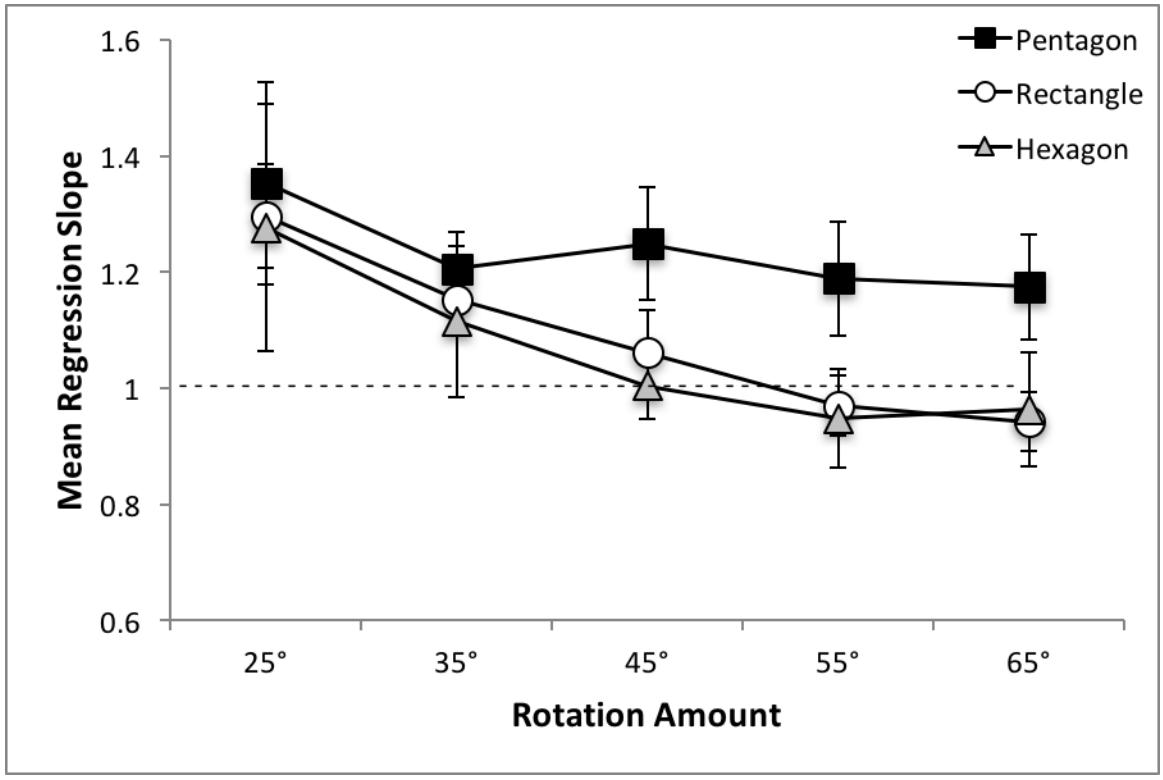


Figure 26. Mean regression slopes in Experiment 6 for generalized cone objects with different cross sections, pentagon, rectangle, and hexagon, displayed using the combination of stereomotion and monocular SFM, plotted as a function of rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Mauclhy's test of sphericity for regression intercepts showed that the sphericity assumption was violated ($\chi^2(9) = 26.23$, $p < 0.01$). With Greenhouse-Geisser correction, ANOVA on regression intercepts showed that there was a significant main effect of rotation amount ($F(2.53,68.25) = 23.20$, $p < 0.001$, $\eta_p^2 = 0.46$). There was neither a significant main effect of objects' cross section shape ($p > 0.5$) or a significant interaction effect between the two factors ($p > 0.1$). Post-hoc LSD showed that intercepts at 25° were significantly smaller than those in other conditions ($p < 0.05$). 35° rotation also had significantly smaller intercepts than 45°, 55°, and 65° of rotation ($p < 0.01$). Intercepts in 45° condition were also smaller than 55° and 65° rotations ($p < 0.01$), and there was no significant difference between 55° and 65° conditions.

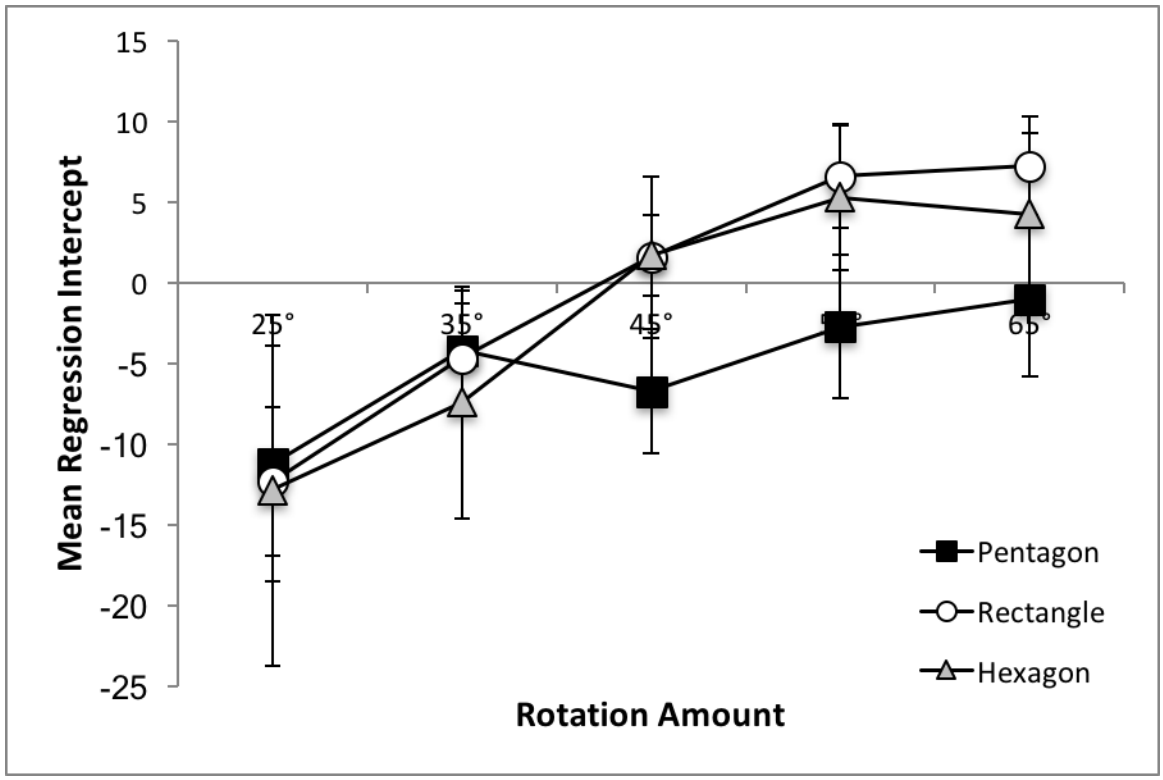


Figure 27. Mean regression intercepts in Experiment 6 for generalized cone objects with different cross sections, pentagon, rectangle, and hexagon, displayed using the combination of stereomotion and monocular SFM, plotted as a function of rotation amount. Error bars represent

95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Subsequently, I examined the means of regression slopes and intercepts with 95% confidence intervals corrected for within-subject designs. As shown in Figure 26 and Figure 27, regression slopes for all three shapes started at above 1 and intercepts below 0 with 25° of rotation. With an increasing rotation, at 35°, slopes for objects with hexagonal cross sections became indistinguishable from 1. At 45° of rotation and beyond, slopes for hexagonal and rectangular objects became indistinguishable from 1. Regression intercepts for the hexagonal objects remained relatively stable at around 0 with further increase of rotation, whereas intercepts for the rectangular objects went beyond 0. This suggested a positive bias as observed in Experiment 2. Slopes for pentagons remained relatively steady and did not became 1 even with 65° of rotation, although their intercepts did become indistinguishable from 0 at 55°. This aspect of the results was something that ANOVA failed to reflect because, in testing only differences between conditions, the between-subject variability was included. This analysis confirmed results from Experiments 2 to 4, where large continuous perspective change enabled metric slant perception for rectangles and hexagons, but not for pentagons, where the difference among the three different shapes was a function of the presence or absence of symmetry.

Finally, regression r^2 was tested. Again, the sphericity assumption was violated ($\chi^2(9) = 18.17$, $p < 0.05$). With Greenhouse-Geisser correction, ANOVA did not show a significant main effect of rotation ($p > 0.2$) or a significant interaction effect between rotation and shape ($p > 0.7$). However, there was a significant main effect of shape ($F(2,27) = 3.65$, $p < 0.05$, $\eta_p^2 = 0.21$). Post-hoc LSD correction should that pentagonal surface had r^2 that were significantly smaller than the rectangular surface ($p < 0.05$). As shown in Figure 28, all three shapes had relatively stable r^2,

and the rectangular surface had on average larger regression r^2 than the pentagonal surface, while the hexagonal surface produced r^2 between the two. Because the pentagonal surface had, on average, greater-than-one regression slopes, small r^2 cannot be due to small regression slopes. Instead, this suggested that in addition to not being accurate, slant judgments of generalized cone with pentagonal cross sections were also variable and less consistent.

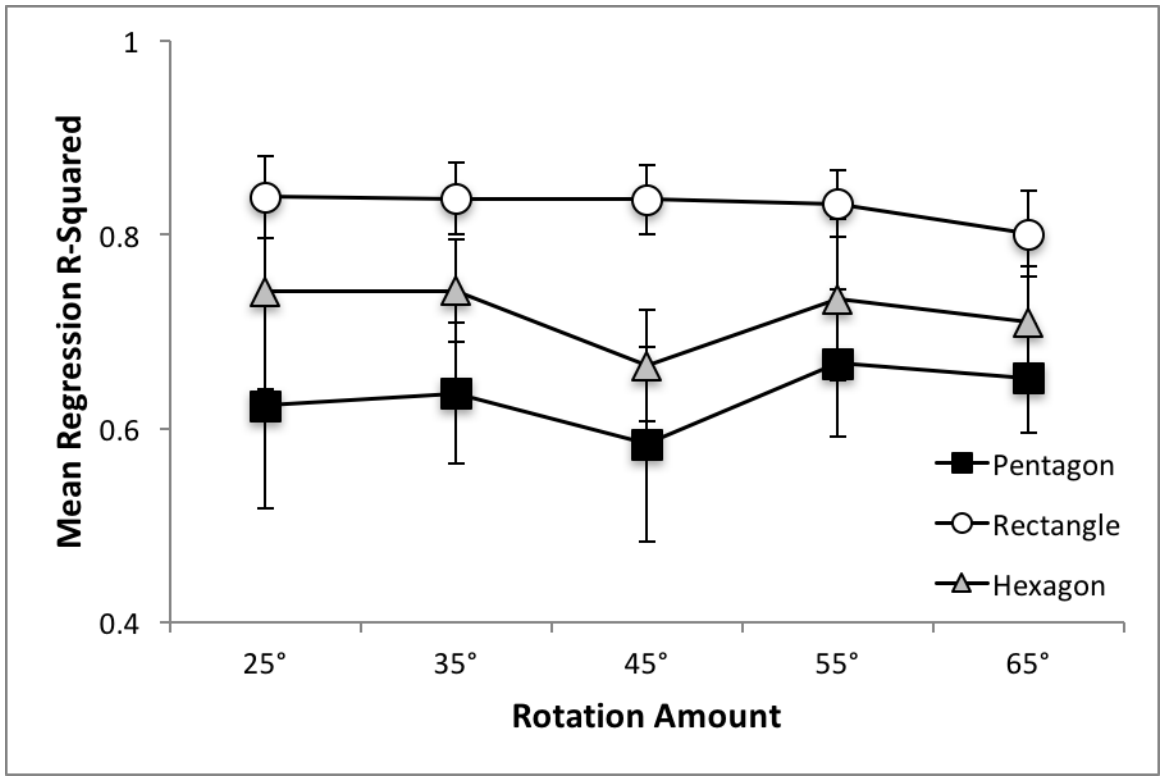


Figure 28. Mean regression r^2 in Experiment 6 for generalized cone objects with different cross sections, pentagon, rectangle, and hexagon, displayed using the combination of stereomotion and monocular SFM, plotted as a function of rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Results from Experiment 6 confirmed several claims in previous experiments. First and foremost, perception of metric slant could be achieved with large continuous perspective change.

This was confirmed with results from the generalized cone objects with rectangular and hexagonal cross sections, similar to the shapes of surfaces used in Experiments 2 and 4. In addition, this also showed that the non-coplanar points, introduced as cuboids and tetrahedrons in previous experiments, could be accomplished using 3D polyhedrons. Third, this confirmed the role of a symmetry axis in specifying the direction of geographical slant. As results from generalized cones with pentagonal cross sections suggested, judging slants that were embedded in 3D polyhedrons was still poor given asymmetric cross sections.

Chapter III: The Model

In Chapter II, I demonstrated in a series of experiments that the bootstrap process could be effectively applied to 3D slant perception, where large continuous perspective change could enable observers in recovering accurate geographical slant. During this process, I discovered that the original right-angle solution to the bootstrap process was not suitable for 3D slant perception and indicated an alternative solution that only used two equidistant points. In Chapter III, I will formalize this alternative solution by presenting a stratified slant recovery process that will eventually be used to simulate behavioral results from Chapter II. In addition, I will also use simulation results to substantiate the claims regarding the role of symmetry in perceiving 3D geographical slants.

Stratified Slant Recovery Process

In this section, I will first illustrate the steps through which one could use the SFM algorithms and the bootstrap process to produce slant estimation. I will first present an implementation of an existing structure-from-motion (SFM) algorithm, which was adopted from Lind (1996). This implementation takes two frames of 2D projected x-y coordinates of a 3D object moving in depth as input and produces a relief depth map of the object and other motion parameter estimates. In the second part, an implementation of the bootstrap process that utilizes the tracking of two equidistant points as initially revealed at the end of Experiment 3 is illustrated. Finally, we present a way through which the resulting scaling factor could be used to produce slant estimation. Throughout this Chapter, I use lower case letters, x, y, and z to denote image coordinates or relief depth (as with z), and capital letters, X, Y, and Z to denote physical coordinates.

Stage 1. Recovery of Relief Depth Map

There are many existing models that can recover relief depth based on two-frame motion (e.g. Koenderink & van Doorn, 1991; Shapiro, Zisserman, & Brady, 1995; Lind, 1996). Because SFM processes have been described and discussed at length in previous studies, we now briefly present a simplified version of the method proposed in Lind (1996). In the case of orthographic projection for SFM analyses, we decompose relative motion between an observer and an object in terms of a rotation of the image plane and a rotation around a unit axis that is in the image plane as shown in Figure 29 (Todd & Bressan, 1990; Koenderink & van Doorn, 1991; Shapiro et al, 1995). Lind (1996) denoted the rotational speed of the image plane as $\dot{q}$ and the rotation around the axis in the plane as $\dot{\alpha}$. In addition, with scaled orthographic (or weak perspective) projection, there is a uniform shrinking or expansion of the 2D image, as a function of the speed in depth, $\dot{z}$. The unit axis that is in the image plane has a direction of δ, relative to the positive x-axis.

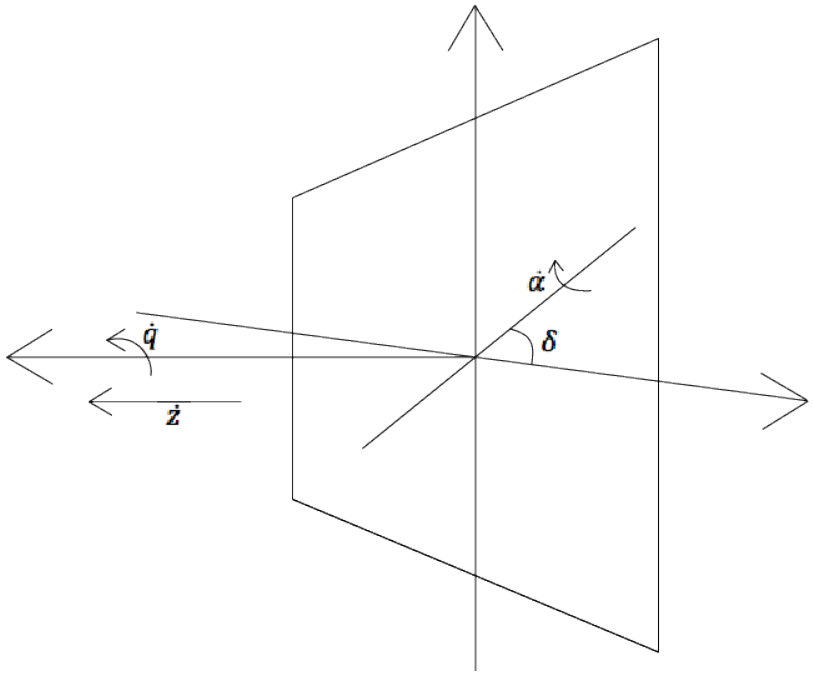


Figure 29. A schematic demonstration of the motion decomposition in Lind (1996). $\dot{q}$ is the angular velocity of the image plane, $\dot{z}$ is the speed of translation in depth yielding a uniform shrinking or expansion of the image plane, δ is the direction of the unit axis in the image plane relative to the positive x-axis, and $\dot{\alpha}$ corresponds to the angular velocity of the rotation around the unit axis. The latter is the essential SFM transformation.

Given such decomposition in a scaled orthographic SFM analysis, one can produce estimates of the unknown parameters $\dot{q}$, $\dot{z}$, and δ with two 2D images based on the viewing of a moving and rigid 3D object. However, the $\dot{\alpha}$ parameter cannot be estimated due to the lack of degrees of freedom in the system. Instead, the depth within an unknown scaling factor ($\dot{\alpha}Z$) is estimated for each image location where Z denotes the distance to this 3D texture element and $\dot{\alpha}$ is the unknown scaling factor. Thus, $\dot{\alpha}$ is the unknown scaling factor in the bootstrap analysis and $\dot{\alpha}Z$ is the relief depth map produced by the SFM analysis.

By aligning the y-axis of a coordinate system in the image plane with the direction δ, the following equations can be obtained:

$$\dot{x} \approx -\dot{\alpha}Z + x\dot{z} - y\dot{q} \qquad \text{Equation 1}$$

$$\dot{y} \approx y\dot{z} + x\dot{q} \qquad \text{Equation 2}$$

where x and y denote the position in the image plane of a specific 3D texture element, and $\dot{x}$ and $\dot{y}$ are the velocity of this texture element across a two-frame apparent motion. Using Equation 2, $\dot{z}_{est}$ and $\dot{q}_{est}$ can be obtained through a linear process, which can then be combined with Equation 1 to produce an $\dot{\alpha}$ scaled depth map:

$$\dot{\alpha}Z \approx -\dot{x} + x\dot{z}_{est} - y\dot{q}_{est} \qquad \text{Equation 3}$$

For more detailed discussions on this topic see Lind (1996).

Stage 2. The Bootstrap Process

At this stage, we use the relief depth map, $\dot{\alpha}Z$, and information provided by the two-frame motion to produce an estimate of $\dot{\alpha}$ that would subsequently be used to produce an estimate of slant. The solution to the bootstrap process presented at this stage is a formalization of the alternative solution presented at the end of Experiment 3. This process starts with the identification of two equidistant points at t_0, P_{1t_0}, $(x_{1t_0}, y_{1t_0}, z_{1t_0})$, and P_{2t_0}, $(x_{2t_0}, y_{2t_0}, z_{2t_0})$. By definition, the length of the line formed between P_1 and P_2 is simply the distance between two points in the image plane (Figure 30):

$$L = \sqrt{\left(x_{1t_0} - x_{2t_0}\right)^2 + \left(y_{1t_0} - y_{2t_0}\right)^2} \qquad \text{Equation 4}$$

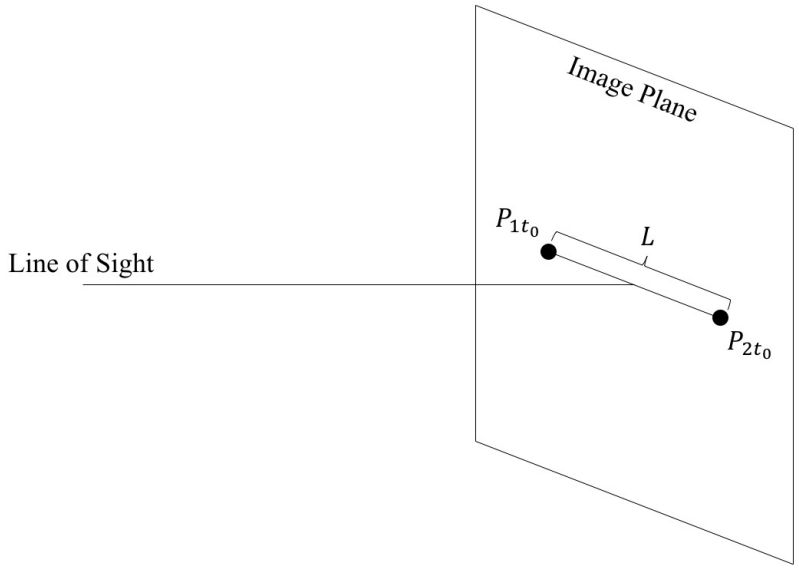


Figure 30. The equidistant point setup at time t_0. P_{1t_0} and P_{2t_0} are equidistant from the observer.

With relative motion between the observer and the object, e.g. motion from rotation around an axis in the x-y plane, P_1 and P_2 will no longer be equidistant or be in the same frontoparallel plane. Let $l(t)$ be the line's orthographically projected 2D length in the image plane at any given time t:

$$l(t) = \sqrt{(x_{1t} - x_{2t})^2 + (y_{1t} - y_{2t})^2} \qquad \text{Equation 5}$$

Then the change in l over time, dt, can be expressed as:

$$\dot{l} = \frac{d(l(t))}{dt} \qquad \text{Equation 6}$$

The angular velocity of rotation at any given time can be expressed as

$$\dot{\alpha}(t) = \frac{\dot{l}}{\sqrt{L^2 - l(t)^2}} \qquad \text{Equation 7}$$

Substituting Equation 4 to Equation 6 to Equation 7 enables one to derive an $\dot{\alpha}$ estimate using only x and y coordinates of the two equidistant points. When noise is added to the system, measuring the projected length of the line could be variable. However, with an increasing amount of perspective change, the projected length would become increasingly smaller, and therefore is less affected by the addition of noise. Therefore, with a larger amount of rotation, $\dot{\alpha}$ estimations should also become more accurate.[1]

Because of the unknown scaling factor, it is difficult for the observer to discern whether there has been a sufficiently large continuous perspective change (just as I had discussed in the case of the right-angle solution). For the right-angle solution, Lind et al. (2014) argued that the observers would be informed when the line of sight bisects the right angle as symmetry and angle bisection are available in relief space. Naturally, bisection of a right angle corresponds to 45° of continuous perspective change. For the equidistant point solution, observers can evaluate the amount of rotation by comparing the change in length of the 3D relief distance between the two initially equidistant points. Norman, Todd, Perotti, and Tittle (1996) found that the threshold for comparing 3D lengths that had different orientation in space relative to the frontoparallel plane was between approximately 19% and 26%. A 20% difference in the 3D line length would correspond to 35.90° of rotation (i.e. change in tilt) whereas a 25% difference would correspond to 42.27° of rotation. This range of continuous perspective change is more flexible than the

---

[1] The current implementation does not take motion in depth into consideration, i.e. $\dot{z} = 0$ *is presumed*. When there is motion-in-depth, the length variables can simply be scaled by the recovered $\dot{z}_{est}$ from the SFM algorithm accordingly.

previous 45° criterion. In fact, depending on the experiments, veridical performance has been found empirically at either 35° or 45° of continuous perspective change (e.g. for the rectangle with cuboids stimuli in Experiment 2, performance became veridical at 35° of rotation).

Computationally, such change can be captured using the angle formed by the line between the two initially equidistant points at time t and the projection of that line in a frontoparallel plane. Because the length of the line should remain the same over rotation, using the projected 2D length of the line at time t and its original length we can calculate the cosine of the angle, θ, formed between the two lines (Figure 31):

$$\cos\theta = \frac{l(t)}{L} \qquad \text{Equation 8}$$

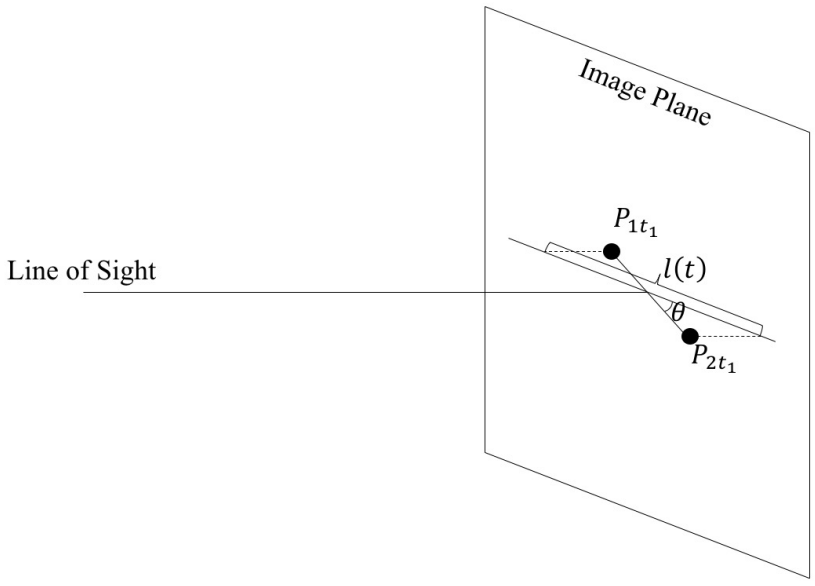


Figure 31. The equidistant point setup at time t_1.

θ in this case directly measures the amount of rotation that has occurred from the first frame. To provide a measure of quality that reflects the usefulness of rotation, I can use the following:

$$qualtiy = 4(\cos\theta - \cos^2\theta) \qquad \text{Equation 9}$$

This measure of quality was computed so that it monotonically increased for θ ranging between 0° and 60°. Based on the criterion angles established earlier, we can compute the corresponding quality measure to evaluate the amount of rotation that has occurred.

Stage 3. Slant Estimation

Given the relief depth map from Stage 1 and the recovered scaling factor from Stage 2, the final step is to recover the perceived slant. Based on discussions in the previous section, we adopted a geographical frame of reference, where slant is defined as the angle between the surface normal and the normal of the ground surface, i.e. unit vector [0,1,0], and tilt is defined as angle between the projection of the surface normal on the image plane (xz-plane) and the unit vector [0,0,1] (Figure 4). For a planar relief surface centered at the origin, its surface equation can be expressed as the following:

$$z = \dot{\alpha}Z = ax + by \qquad \text{Equation 10}$$

where z is the relief depth, which is equivalent to its actual depth, Z, scaled by the scaling factor $\dot{\alpha}$. Since we have already obtained z and $\dot{\alpha}$ in the first two stages, we can again use linear processes to acquire coefficients a and b. Subsequently, the plane equation can be expressed as:

$$ax + by - \dot{\alpha}Z = 0 \qquad \text{Equation 11}$$

The surface normal is therefore $(a, b, -\dot{\alpha})$. We can then find the angle between the surface normal and unit vector [0,1,0] using cross and dot products:

$$\sin(\sigma) = \frac{\sqrt{\dot{\alpha}^2 + a^2}}{\sqrt{a^2 + b^2 + \dot{\alpha}^2}} \qquad \text{Equation 12}$$

$$\cos(\sigma) = \frac{b}{\sqrt{a^2 + b^2 + \dot{\alpha}^2}} \qquad \text{Equation 13}$$

Dividing Equation 12 by Equation 13 yields a simpler expression of σ:

$$\tan(\sigma) = \frac{\sqrt{\dot{\alpha}^2 + a^2}}{b} \qquad \text{Equation 14}$$

Because $\dot{\alpha}$ is known from the bootstrap model, slant can be estimated.

Additionally, results from Experiment 5 showed that when the direction of slant is not aligned with the direction of the symmetry axis of an object, perception of slant tended to reflect that along the symmetry axis. To address the issue of slant direction, I will use a second method to derive slant estimates. First, for symmetrical objects, I assume that the observers can correctly identify the symmetry axis of an object. There are many studies that model the perception of symmetry and skewed symmetry axes (see e.g. Li, Sawada, Shi, Kwon, & Pizlo, 2011). With the symmetry axis on the object, along with the object's normal, I can find a plane that passes through both vectors. I computed slant to be the angle between the normal of that surface and the unit vector [0,1,0]. For asymmetrical objects, namely the pentagon surface with tetrahedrons used in Experiment 4, a random point would be chosen along the edge of the surface in our simulation and be connected to the pentagon's top vertex as the direction of slant to reflect the uncertainty of the direction of slant. Therefore, this could be a test of whether in fact observers exhibited random slant directions in their judgments when the surface lacks symmetry.

Model Simulation

I now look at how model predictions compared to actual human performance. I performed simulations of four experiments with three different types of objects from Experiments 2 to 5. The first type of object that I simulated was a rectangular surface with nine cuboids on top in a grid, as used in Experiment 2. This is to test if the proposed stratified process

could predict human performance, i.e. whether sufficiently large continuous perspective change would allow accurate slant estimation. This object contained the most structure (right angles and symmetry) compared to the others to be used, so it provided a baseline test of the effectiveness of the algorithms. The second type of object used in simulations was hexagonal with tetrahedrons on top at random locations, as in Experiment 3. Additionally, I used this object with changes in the orientation of the symmetry axis, which was then used to derive alternative slant estimates, as in Experiment 5. These two simulations tested the claim that observers used the direction of the symmetry axis to determine the slant direction. Finally, I used an asymmetric pentagonal surface with randomly placed tetrahedrons to simulate the results in Experiment 4. In this simulation, slant was computed using a random reference line on the object as the direction of slant. This simulation tested the claim that with a lack of symmetry axis, observers randomly chose the direction of slant.

I did not simulate performance with strictly planar surfaces because the lack of non-coplanar points would prevent the SFM process from recovering the unknown motion parameters in the beginning of the stratified process. Specifically, the planar surface would produce one less linearly independent equation than the number of unknown motion parameters for the system to solve. Recall from Stage 1 of the stratified process, this means that either $\dot{z}_{est}$ or $\dot{q}_{est}$ was still remain unknown after the SFM process. However, the current experimental setup lacks $\dot{z}$ and $\dot{q}$ motion components, allowing the recovered relief depth map to be usable. This does not apply to the human visual system and the resulting depth map cannot be confidently used for the subsequent bootstrap process. To explore the effects of the lack of non-coplanar points on the bootstrap process, we incorporated motion-in-depth, i.e. $\dot{z}$, in a simulation, comparing the effectiveness of the bootstrap process in recovering the unknown scaling factor with a

rectangular planar surface and the same surface with additional cuboids. We simulated a continuous rotation of 90°, combined with a translation along the z-axis of 4 cm during the rotation. I identified 50 pairs of equidistant points and computed the estimated scaling factor for each pair across rotation. The final result at each tilt was the mean of those pairs. Figure 32 shows the simulation results. The bootstrap process could successfully recover the unknown scaling factor for the surface with additional 3D structures in the presence of motion-in-depth, but not for the strictly planar surface.

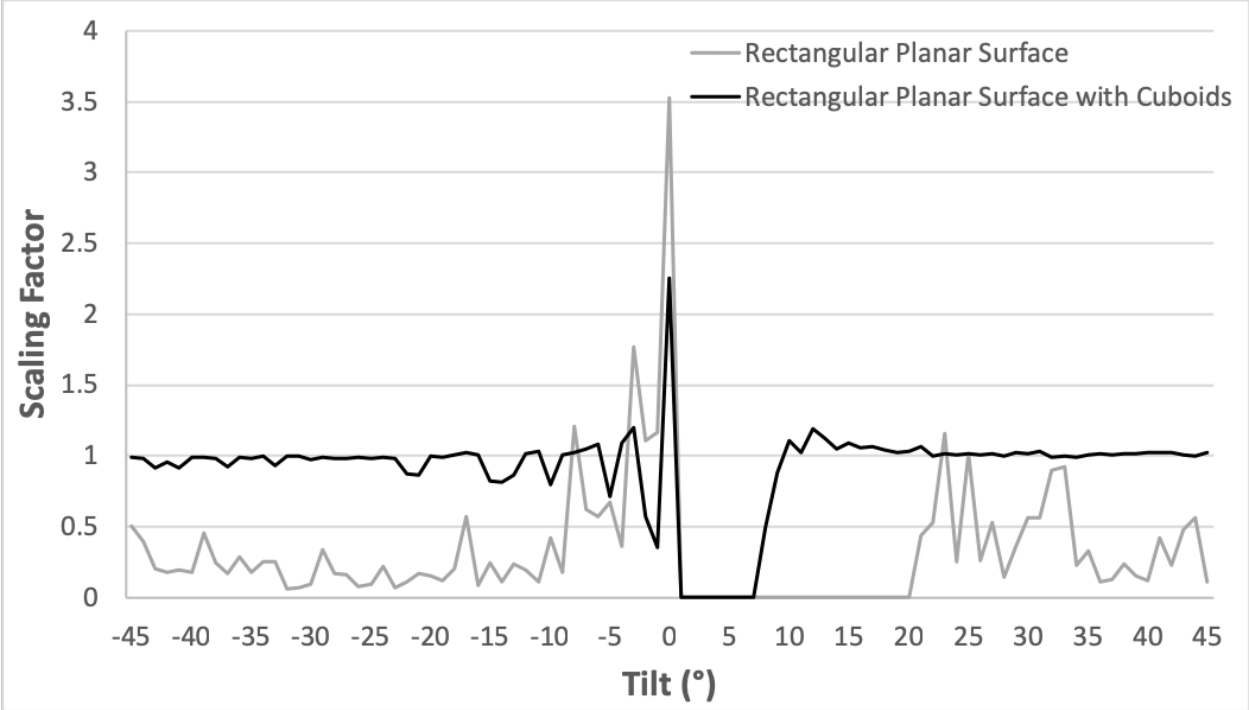


Figure 32. Recovered scaling factors through the bootstrap process for a strictly rectangular planar surface and the same surface with nine cuboids as a function of change in tilt. The correct scaling factor is 1.

Experimental Stimuli and Setup

The three different types of objects used both in human experiments and in simulations are illustrated in Figure 33, including the rectangular surface with nine cuboids located in a 3-by-3 grid (Experiment 2), the hexagonal surface with nine tetrahedrons at random locations (Experiment 3), and the asymmetrical pentagonal surface with nine tetrahedrons at random locations (Experiment 4), together with the manipulation of the orientation of the symmetry axis of the hexagonal surface where I rotated the hexagonal surface around its surface normal by 15° (Experiment 5). As described in previous experiments, all surfaces had the same width, 10 cm, and heights, 8 cm, 10cm, or 12 cm. The cuboids and tetrahedrons had heights of 0.55 cm and base length of 1 cm.

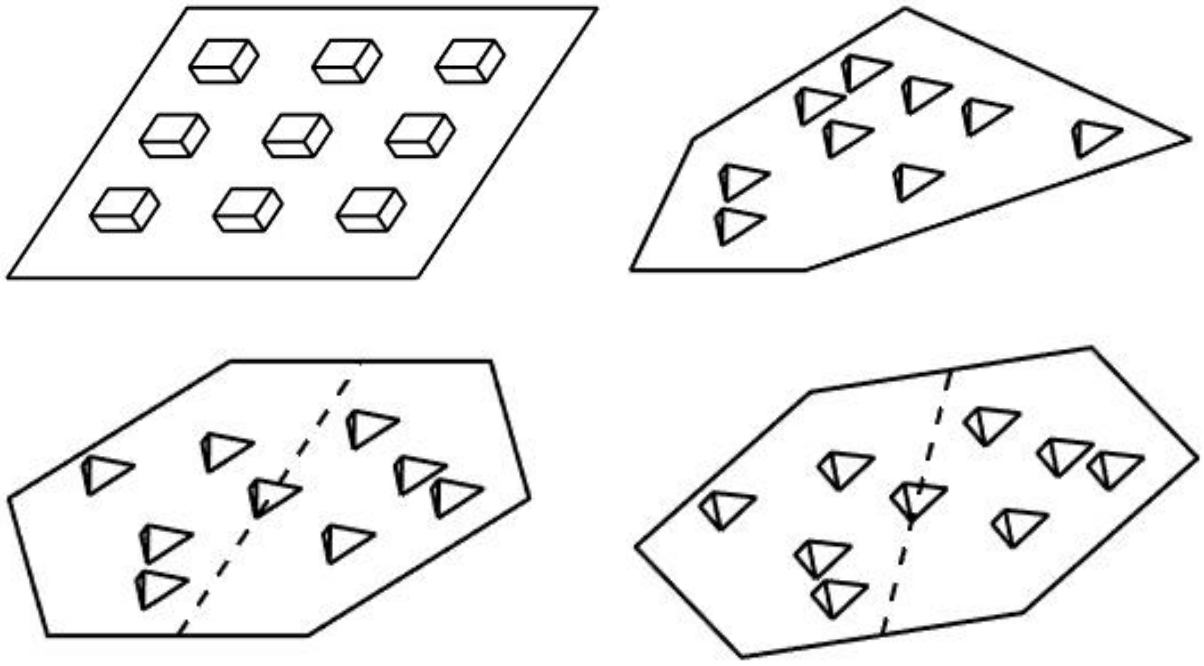


Figure 33. Schematic illustration of surfaces used in experimental displays and model simulations: rectangular surface with nine cuboids in a three-by-three grid (top left), asymmetrical pentagonal surface with nine randomly located tetrahedrons (top right),

hexagonal surface with randomly located tetrahedrons (bottom left), and the same hexagonal surface with the symmetry axis rotated around the surface normal by 15° (bottom right). For the hexagonal objects, dashed lines represent the direction of symmetry axis. All objects presented have a slant of 45° and a tilt of -32.5°.

Simulation followed the same empirical setup as the experiments. Specifically, there were 5 different rotation amounts, from 25° to 65° with a 10° increment. Objects were presented using half rotation. For instance, with a 25° rotation, the object, started from facing the observer, at a tilt of 0°, first rotated to half of the amount of rotation at 12.5° to one side, rotated back to the starting orientation, and then rotated to 12.5° of the other side. Within each rotation amount, there were a total of 24 different slant angles, from 27° to 73° with a 2° increment. Finally, for three consecutive slant angles, the height of the surface pseudo-randomly varied among 8 cm, 10 cm, and 12 cm. Therefore, for each simulated shape, there was a total of 120 trials (5 rotation amounts times 24 different slant angles). To match sample size of different experiments, I simulated 12 participants' data using the rectangular strictly planar surface experiment, 10 using the rectangular surface with cuboids experiment, 11 using the hexagonal surface with tetrahedrons experiment, and 10 using the asymmetric pentagonal surface with tetrahedrons experiment. Stimuli were presented using the monocular SFM information in the actual experiment. The objects were consisted of random dots, located 9 cm behind and screen. They were back projected onto the screen from a projection point 76.2 cm in front of the screen, yielding the viewing distance to be 85.2 cm.

Simulation Methods

Each simulated object consisted of the same random texture points as used in the corresponding human experiment, with a texture density of 26.67 points per cm^2. I first constructed the objects using these points, rotated the objects, and projected points on the object to the screen, generating x and y coordinates for each point. Coordinates were then scaled by the viewing distance, which was 85.5 cm. Gaussian noise with a standard deviation of 0.002 was introduced to the coordinates to reflect random error associated with motion measurement. Each frame had different noise seed.

I generated inputs to the model for each trial. Within each trial, a set of 50 samples each consisting of two equidistant points found on the surface were identified. The model then tracked these samples across the entire range of rotation, producing a quality measure, $\dot{\alpha}$, and the surface plane coefficients. The model first checked among the samples whether the respective quality measures satisfied a criterion. The quality criterion was set based on a 19% 3D distance discrimination threshold (Norman et al., 1996). I first identify samples that had quality measure exceeds the criterion. If such samples exist, I used the median $\dot{\alpha}$ estimates at the tilt where the quality measure was above criterion for those samples, and the final $\dot{\alpha}$ estimate was the mean of the medians. Alternatively, if no sample satisfied the quality criterion, I simply used the mean of the medians of all samples across all tilts as the final $\dot{\alpha}$ estimation.

Furthermore, as numerous studies have suggested, human observers tended to perceive more upright slants to be closer to the frontoparallel plane (and thus, as more slanted) (Todd, Thaler, & Dijkstra, 2005; Norman et al., 2009; Durgin, Li, & Hajnal, 2010; Saunders & Chen, 2015). In addition, Cherry and Bingham (2018) reported a similar bias towards shallower slants, where observers tended to perceive smaller slants to be closer to the horizontal ground surface.

According to the authors, such a bilateral bias could be attributed to the presence of vertical (e.g. walls) and horizontal (e.g. table) references in the experiment. Similarly, Saunders and Chen (2015) argued that for computer generated displays, presence of the vertical computer screen could generate a frontal bias. To represent these biases, we introduced additional noise to the final $\dot{\alpha}$ estimate in the model that was proportional to the derived slant when no sample had a quality measure that satisfied the criterion. Such noise would make large slants more frontoparallel and small slants more horizontal.

Finally, based on findings from Experiments 4 and 5, the direction of slant could also vary depending on the presence or absence of mirror symmetry in the object. To reflect this issue, for the pentagonal objects, I used the symmetry axis as its slant direction. For the pentagonal objects, I picked a random point along the bottom edge of the object and connected it with its top vertex as its slant direction. I varied the slant direction for pentagonal objects for every trial.

Data Analysis

To analyze and compare model prediction and actual human performance, I used the same analysis protocol as in the Chapter II. Specifically, I used linear regression, regressing predicted slant onto actual slant, and regression slopes and intercepts for each rotation amount and each participant as measures of performance. A veridical judgment of slant within a rotation condition entails a regression slope of 1 and a regression intercept of 0. To compare model prediction with human performance, I used mixed design analysis of variance (ANOVA) on regression slopes, intercepts, and r^2 with one within-subject factor of rotation amount (five

levels) and one between-subject factor of type of results (two levels, human performance or model prediction).

Simulation Results

For the rectangular surface with cuboids, regression slopes and intercepts for human performance and model simulation are shown in Figure 34. Mauchly's test of sphericity showed that the sphericity assumption was violated ($\chi^2(9) = 27.55$, $p = 0.001$). With Greenhouse-Geisser correction, ANOVA on regression slopes showed that there was a significant main effect of rotation amount ($F(2.50,55.00) = 7.00$, $p = 0.001$, $\eta_p^2 = 0.24$). There was no significant effect of the type of results ($p > 0.9$, $\eta_p^2 = 0.00$) or a significant interaction effect between the two ($p > 0.7$, $\eta_p^2 = 0.015$). For regression intercepts, the sphericity assumption was also violated ($\chi^2(9) = 35.65$, $p < 0.001$). ANOVA on regression intercepts showed that there was a significant main effect of rotation ($F(2.40, 52.72) = 7.44$, $p = 0.001$, $\eta_p^2 = 0.25$) after Greenhouse-Geisser correction. There was neither a significant main effect of source of data ($p > 0.2$, $\eta_p^2 = 0.05$) nor a significant interaction between the two factors ($p > 0.1$, $\eta_p^2 = 0.07$). As Figure 34 shows, human performance and model prediction are identical for regression slopes. Although regression intercepts were a little higher for human performance than for model prediction, such a difference failed to become statistically significant.

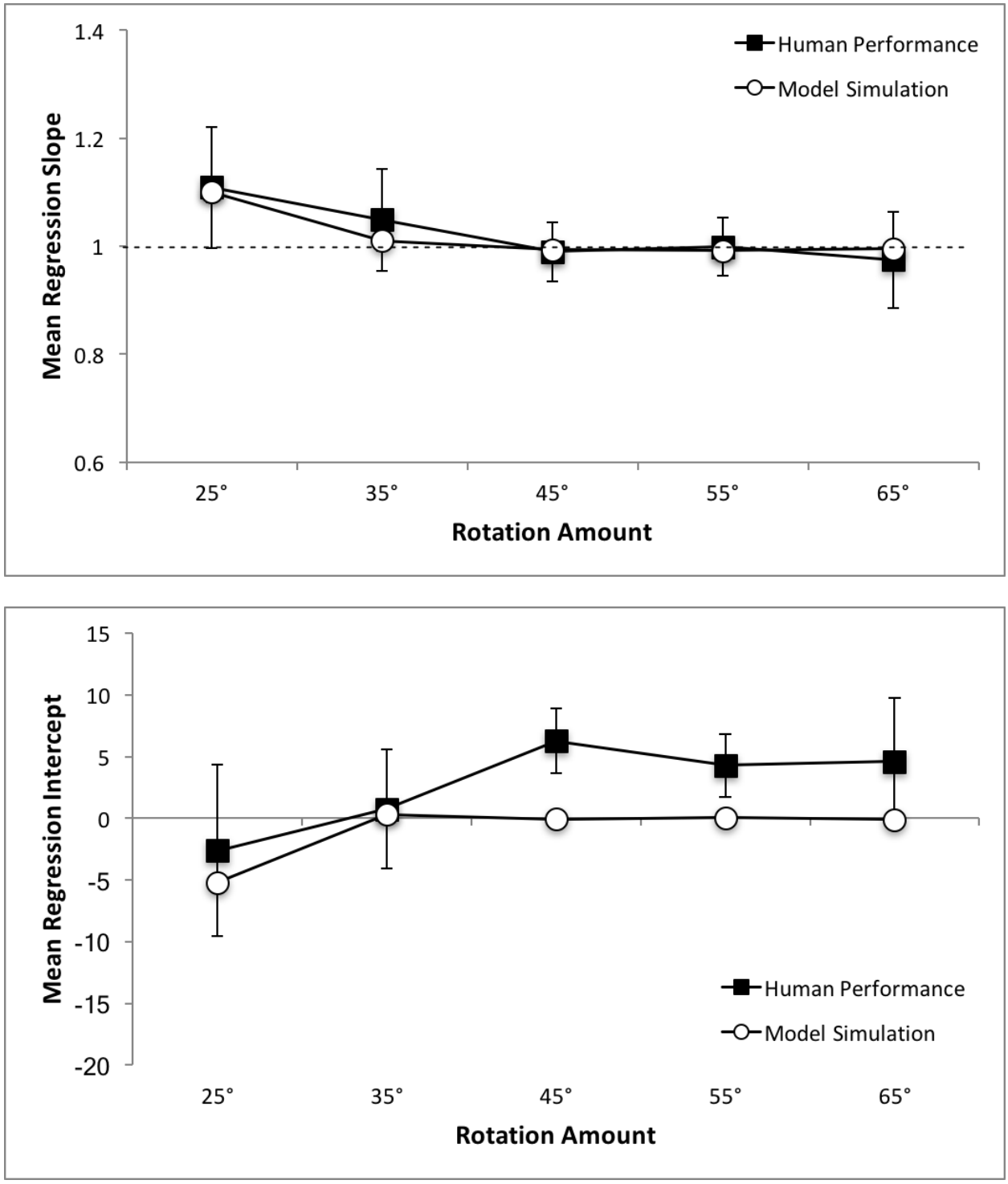


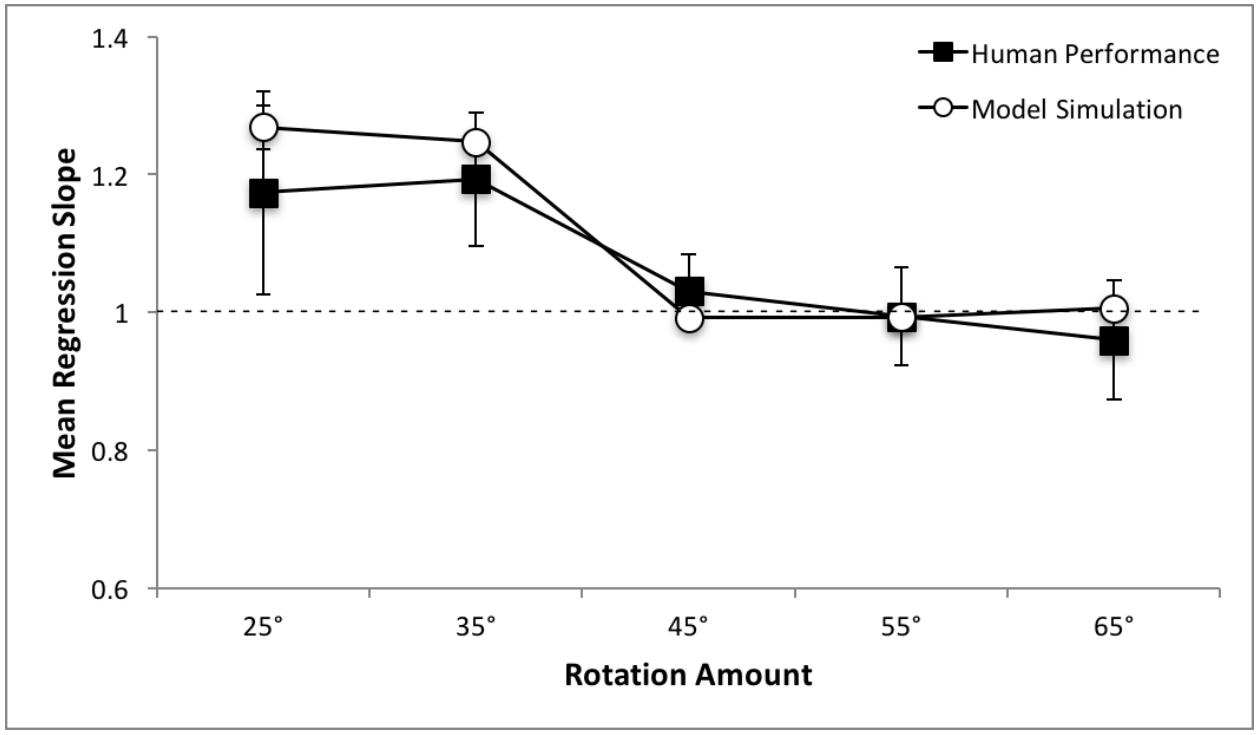
Figure 34. Mean regression slopes (top) and intercepts (bottom) for human performance and model simulation using the rectangular surface with cuboids. Error bars represent 95%

confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Based on comparisons of human and simulation results for the cuboids display, one can see that the equidistant points bootstrap process works well. In particular, slant estimation was inaccurate at 25° of rotation. With 35° rotation, model performance improved a little but did not reach the veridical level. Finally, at 45° and beyond, slant estimations became accurate. Results from this simulation had two implications. First, the stratified process of deriving slant estimates using the equidistant points implementation is suitable for describing performance in 3D slant perception. As the results suggested, model performance did improve with an increasing amount of rotation and remained at a steady level once rotation reached and went beyond 35°. Secondly, the additional noise introduced to the final estimates of the scaling factor when the quality measure did not exceed threshold accurately reflected the tendency of human perception of 3D slant, namely that people tended to overestimate large slant (more upright slants were perceived to be even more upright).

Next, using the same method, I generated model simulations of the hexagonal surface with tetrahedrons at random locations. Figure 35 shows the mean regression slopes and intercepts for human and model performance. Again, we saw identical performance in model and data. For regression slopes, the sphericity assumption was violated ($\chi^2(9) = 30.68$, $p < 0.001$). With Greenhouse-Geisser correction, ANOVA showed that there was a significant main effect of rotation amount ($F(2.35,47.04) = 31.59$, $p < 0.001$, $\eta_p^2 = 0.61$). However, there was neither a main effect of the type of results ($p > 0.5$, $\eta_p^2 = 0.02$) nor a significant interaction effect between the two factors ($p > 0.2$, $\eta_p^2 = 0.064$). For regression intercepts, the sphericity assumption was

violated ($\chi^2(9) = 19.73$, $p < 0.05$). With Greenhouse-Geisser correction, ANOVA showed that there was a significant main effect of rotation amount ($F(2.79,55.80) = 42.01$, $p < 0.001$, $\eta_p^2 = 0.68$). There was neither a main effect of type of results ($p > 0.8$, $\eta_p^2 = 0.002$) or a significant interaction effect between the two factors ($p > 0.3$, $\eta_p^2 = 0.05$). As can be seen from Figure 35, the model predicted the pattern of performance for the hexagonal objects very well. The model successfully predicted veridical performance at 45° of rotation and beyond.



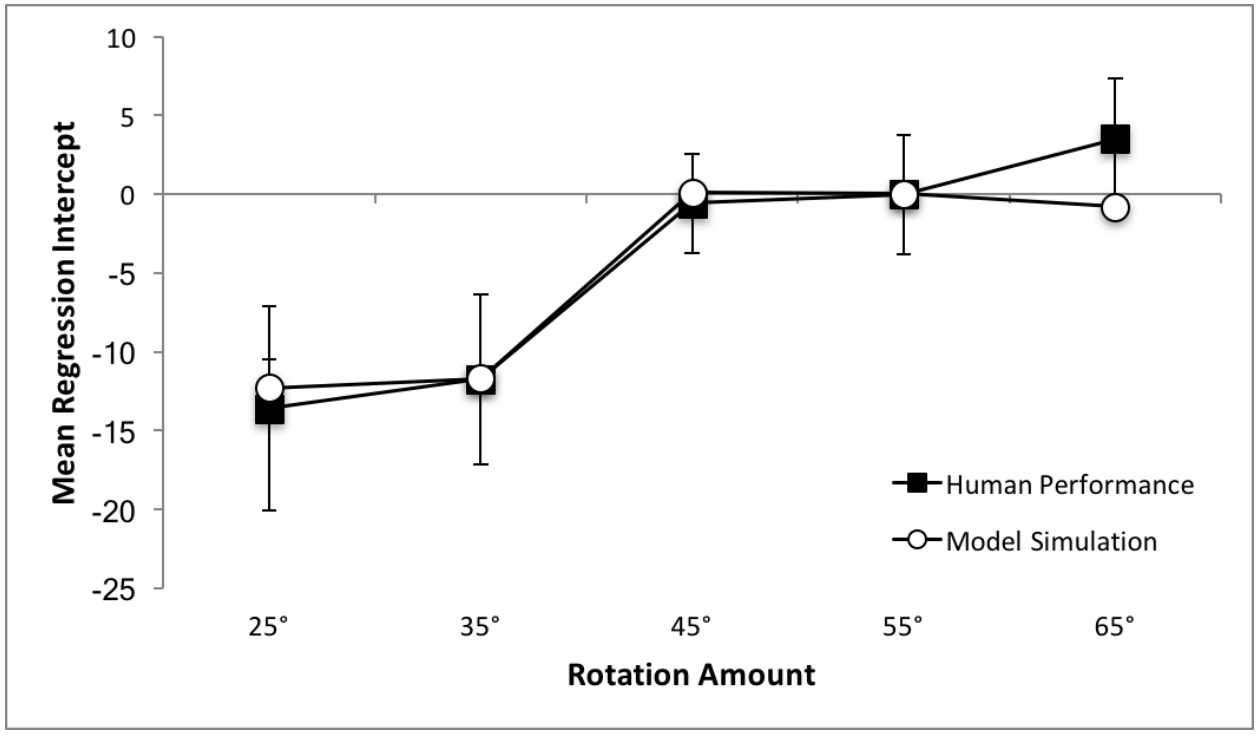
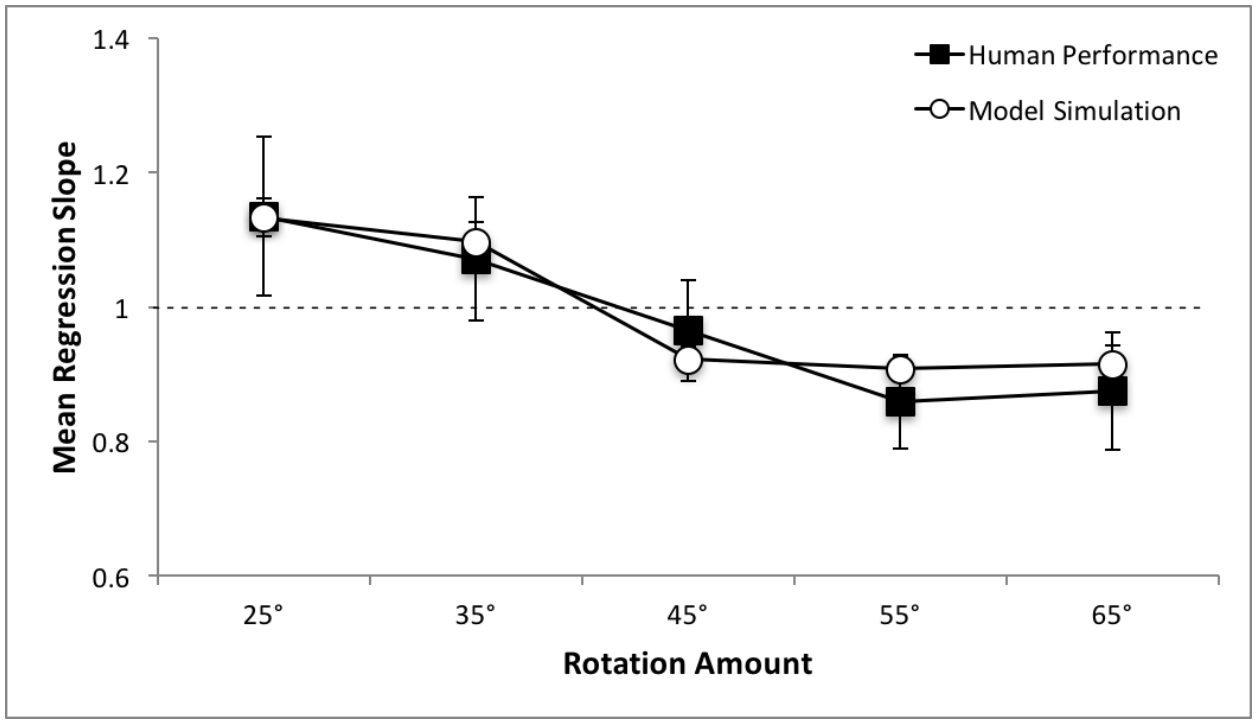


Figure 35. Mean regression slopes (top) and intercepts (bottom) for human performance and model simulation using the hexagonal surface with random tetrahedrons. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

Next, I performed simulations for hexagonal displays with a 15° roll. In my simulation, to explore the role of symmetry, I computed slants based on the direction of the symmetry axis instead of using the surface normal. Figure 36 shows the mean regression slopes and intercepts for human performance and model simulation. Mauchly's test of sphericity showed a violation of the sphericity assumption for regression slopes ($\chi^2(9) = 21.87$, $p = 0.01$). With Greenhouse-Geisser correction, there was a significant main effect of rotation amount ($F(2.63,47.39) = 25.03$, $p < 0.001$, $\eta_p^2 = 0.62$). There was no significant main effect of the source of data ($p > 0.7$, $\eta_p^2 =$

0.007) or a significant interaction effect between the two factors ($p > 0.5$, $\eta_p^2 = 0.041$). For regression intercepts, there was also a significant main effect of rotation amount ($F(4,72) = 28.88$, $p < 0.001$, $\eta_p^2 = 0.62$). Again, there was neither a not a significant main effect of source of data ($p > 0.8$, $\eta_p^2 = 0.005$) or a significant interaction effect ($p > 0.2$, $\eta_p^2 = 0.076$).



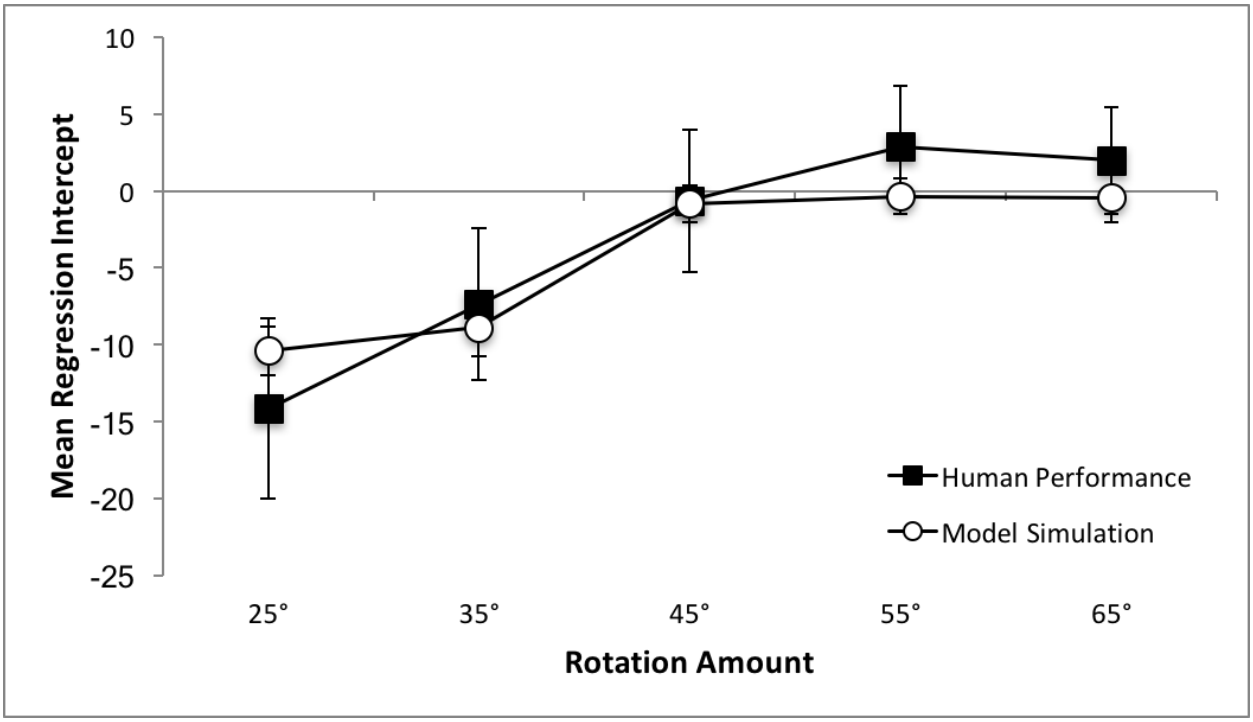


Figure 36. Mean regression slopes (top) and intercepts (bottom) for human performance and model simulation using the hexagonal surface with random tetrahedrons with a 15° rotation of the surface around its normal. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

As shown in Figure 36, model simulation produced the same pattern as human performance both in terms of regression slopes and intercepts. In Experiment 5, I suggested that observers judged slant using the symmetry axis to determine the direction of slant. I subsequently regressed slant judgments onto slants that were computed along the direction of the object's symmetry axis. With this, the dependent measures reached veridical level (i.e. slope of 1 and intercept of 0). In my simulation, after using the recovered relief scaling factor to adjust recovered depth values, I computed slant along the direction of the symmetry axis and regressed slant estimates from the model onto actual (correct) slants. This produced results that are

equivalent to human performance when it is compared to actual (correct) slants. This provided additional support for the claim that human observers did in fact used the object's symmetry axis as the direction of slant and that they could still recover the correct scaling factor even though symmetry was perturbed.

Finally, I simulated results from the pentagonal surface with tetrahedrons. In Experiment 5, I suggested that poor performance in this case was due to the lack of specification of the direction of slant. To simulate such an effect, for each trial, I randomly chose a point along the bottom edge of the pentagonal surface and formed an axis connecting that point to the top vertex of the surface. This yielded poorly constrained and variable estimates of the slant direction. Figure 37 shows the mean regression slopes and intercepts. Again, sphericity assumption was violated for regression slops ($\chi^2(9) = 19.33$, $p < 0.05$). With Greenhouse-Geisser correction, ANOVA showed that there was a significant main effect of rotation ($F(2.68,48.20) = 10.65$, $p < 0.001$, $\eta_p^2 = 0.37$). However, there was neither a significant main effect of source of data ($p > 0.4$, $\eta_p^2 = 0.033$) or a significant interaction effect between the two tested factors ($p > 0.05$, $\eta_p^2 = 0.12$). For regression intercepts, the sphericity assumption was violated ($\chi^2(9) = 21.69$, $p < 0.05$). With Greenhouse-Geisser correction, there was a significant main effect of rotation ($F(2.60,46.81) = 8.77$, $p < 0.001$, $\eta_p^2 = 0.33$). There was no effect of source of data ($p > 0.1$, $\eta_p^2 = 0.08$) but there was a significant interaction effect between the two factors ($F(2.60,46.81) = 3.50$, $p < 0.05$, $\eta_p^2 = 0.16$).

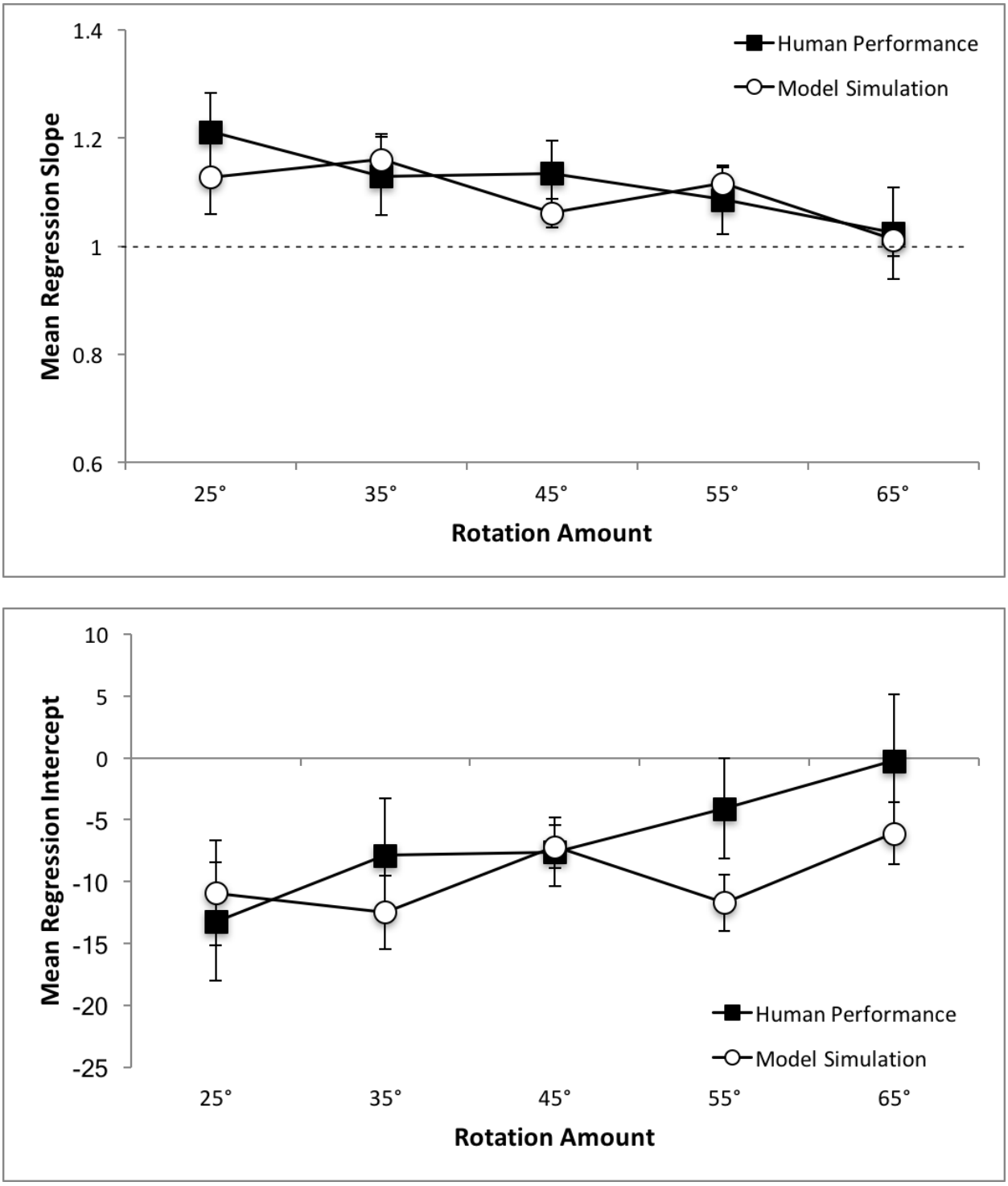


Figure 37. Mean regression slopes (top) and intercepts (bottom) for human performance and model simulation using the pentagonal surface with random tetrahedrons. Error bars represent

95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008).

As seen from Figure 37, regression slopes from model simulations again replicated those from human performance. There was a trend toward veridical performance over increasing amounts of rotation despite the random estimates of the direction of slant. This was true of both regression slopes and intercepts. However, there was a discrepancy between model simulation and human performance in intercepts at 55° of rotation, which contributed to a significant interaction effect. As seen from Figure 37, the model yielded smaller intercepts than human results. Overall, the significance of this simulation is that it provided evidence for arguments made in Experiments 4 and 5 that for the pentagonal surfaces, due to the lack of symmetry, observers chose the direction of slant randomly for each trial. Because the model was able to produce accurate estimates of the unknown scaling factor, this suggested that observers in the pentagonal experiment were able to recover the correct scaling factor as well.

Chapter IV: General Discussion

Human perception, in particular, the perception of 3D objects, is a topic that has captivated the imaginations of, and motived empirical investigations performed by a countless number of philosophers and scientists alike. Coming from a functionalist's perspective, studying the perception of 3D objects is not necessarily about how to translate from basic physical and neural inputs to the ever-elusive perceptual phenomenon. Instead, it is to describe the underlying mechanisms through which human observers interact with their surroundings. The bootstrap process is thus inspired by this latter way of thinking. Observers are constantly moving, producing relative motions that yield continuous perspective changes. Because this latter fact has often been ignored, it is not surprising that numerous studies have reported that 3D shape perception was poor given that they all, to some extent, limited observers' freedom of movement.

This is particular the case in the context of slant perception. There are two predominating lines of investigation in the field of slant perception, one revolves around an egocentric frame of reference and the other, an exocentric frame of reference. For the former, experimental stimuli either were static or entailed little relative motion between target surfaces and the observer. Various types of models have been proposed to account for this kind of slant perception, especially those from a Bayesian cue integration perspective (e.g. van Ee, Adams, & Mamassian, 2003; Knill, 2007). The problem with this approach, as mentioned earlier, is that it does not offer constancy: as soon as the line of sight moves, the perceived slant should change as well. This is rather counterintuitive because, for instance, the perceived slant of the flat ground surface would be constantly changing as the observer moved his/her line of sight. On the other hand, there have been studies that focused on exocentric slant perception, that is, perception of geographical slant,

for instance, studies by Proffitt and colleagues, who primarily studied the effect of psychological states on perceived slant (e.g. how an implied increase in effort caused by wearing a heavy backpack might make an observer judge a hill to be steeper; Bhalla & Proffitt, 1999; Proffitt, Creem, & Zosh. 2001). This line of work, in addition to receiving various criticisms of methodology (e.g. demand characteristics; see Firestone (2013)), has rarely investigated the information to specify the slants in question. The theoretical and empirical endeavors described in this dissertation pursued both the question of information and the promise of constancy in geographical slant, seeking the information content that could aid observers to extrapolate invariant characteristics of slant through transformation, with the help of the bootstrap process.

In Experiment 1, strictly rectangular planar surfaces were used to test if the bootstrap process could be applied to slant perception. Performance in all three visual information conditions failed to improve. This was due to the lack of 3D non-coplanar points on the object, preventing the observers to attain the 3D relief structure required by the bootstrap process in the first place. In contrast, Experiment 2 presented the same rectangular surfaces with nine cuboids on top in a grid as a means to introduce non-coplanar points. Performance improved to become accurate at 35° and remained relatively stable with more rotation. This experiment confirmed that the bootstrap process could be applied to 3D slant perception. More importantly, the contrast between Experiments 1 and 2 was the first set of results indicating that the bootstrap process operates on the level of 3D, instead of 2D retinal images. In addition, results from the stereomotion condition from Experiment 2, where there was a lack of trackable texture elements, either 2D or 3D, suggested that the original formulation of the bootstrap process, namely identifying points that make up right angles and tracking them through rotation, could be reformulated into the tracking of 3D structures.

Given that there are potential repercussions for 3D slant perception related to the original right-angle implementation of the bootstrap process, it is also possible that the right angle itself is not needed for bootstrapping metric slant. Therefore, in Experiment 3, the inherent right angles were eliminated from the stimuli by testing symmetric pentagonal surfaces with randomly located tetrahedrons. Performance still reached veridical level at 45° of rotation for both stereomotion and combined conditions. Most notably, participants could still recover metric slant with large amounts of perspective change in the stereomotion condition, where there was a lack of trackable texture elements and inherent right angles. Such results confirmed not only that the bootstrap process operates in a 3D relief space, but also that the original right-angle solution to the bootstrap process was not suitable to account for metric slant perception. Then, I proposed an alternative solution, using only a subset of the original right-angle 3D structure, namely two equidistant points, to perform the bootstrap process.

In the subsequent experiments, I investigated the sole of symmetry on geographical slant perception, given the intuition that the direction of slant could be ambiguous in my experimental paradigm. In Experiment 4, the symmetry aspect of the stimuli was perturbed with asymmetric pentagonal surfaces and randomly located tetrahedrons. Performance failed to improve with large continuous perspective change. I suggested that symmetry merely helped to specify the direction of slant, rather than being an essential component to the bootstrap process. Stimuli used thus far all had symmetry axes that coincided with the direction of slant. With the lack of symmetry, the direction of slant could become ambiguous, or less well or precisely perceived. This implied that observers could still recover the correct scaling factor, but only failed to be able to use it to recover the correct slant. In Experiment 5, this was tested with the same hexagonal objects used in Experiment 3 but with a different roll so that the slant specified by

their symmetry axes was different from the actual slant. Results from this experiment confirmed that observers did judge slant based on the direction of the symmetry axis.

Finally, results from all of these experiments were replicated using generalized cone objects in Experiment 6 to make the experimental stimuli more representative of what people would see in a natural environment. In addition, this was also used to test if the "bumps" on top of various 2D shapes were the essential component in providing non-coplanar points or flat surfaces and whether other forms of 3D structure could also achieve the same goal. With generalized cones that had rectangular, pentagonal, and hexagonal cross sections, performance was exactly the same as that in previous equivalent experiments: performance became veridical either at 35° or 45° of rotation for generalized cone objects with hexagonal or a rectangular cross sections, respectively, while performance remained relatively poor for those with a pentagonal cross section.

Based on the empirical findings, I presented a new implementation of the bootstrap process based on identifying and tracking two equidistant points across large rotation. I presented a stratified process, from using 2D image inputs to obtain relief structures through SFM processes, to bootstrapping the correct scaling factor through the equidistant point method, and then, using the scaling factor to produce a slant estimation. Model simulations based on Experiments 2 to 4 were presented, all of which produced predictions that matched human performance. Results from model simulation served several different purposes. First and foremost, I tested if the equidistant points method is suitable for describing bootstrapping 3D slant perception, as with the results from simulation of the rectangular surface with cuboids from Experiment 2. Second, I aimed to test if eliminating right angles in the stimuli would affect performance, as in Experiment 3, and it did not. Third, given that the bootstrap process could

successfully recover the correct scaling factor at around 45° of perspective change, the model used to confirm if observers truly did use random slant direction with the asymmetrical pentagon objects. In fact, model simulation results were almost identical to the actual performance, further substantiating the fact that without symmetry, observers could still recover the correct scaling factor, but they then used it to recover the wrong slant. This was also further shown with simulation of the hexagonal objects with a different roll, as simulation results were identical to human performance if deriving the slant in simulation was based on the symmetry axis.

This study presents an alternative way for obtaining metric structure based on relief structure. Traditionally, SFM algorithms have been widely used in the field of computer vision to reconstruct 3D objects or scenes based on 2D image inputs (e.g. Huang & Netravali, 1994; Nistér, 2005). Because the recovered depth from these algorithms is subject to relief scaling, researchers have tried to find methods by which metric depth can be obtained. A common approach is to use stereoscopic calibration, where two or more cameras are used to recover depth of a common scene and, given known camera setup, the correct depth structure can be recovered (e.g. Zisserman, Beardsley, & Reid, 1995). Research from this dissertation presents an alternative solution to this problem.

In addition, this study also yields an interesting implication regarding the debate on optical and geographical slant. There has been a long history of studies that defined slant as the angle formed between a surface normal and the line of sight (e.g. Gibson, 1950; Gibson & Cornsweet, 1952; Stevens, 1983a; Todd & Perotti, 1999; Saunders & Knill, 2001; Norman et al., 2006; Sawada & Pizlo, 2008). However, as Sedgwick and Levy (1985) showed, observers are not adept at judging the slant of a surface based on the line of sight. They argued that it would be more efficient to define slant using an environmental frame of reference, relative to the direction

of gravity. They called this geographical slant. Cherry and Bingham (2018) explicitly manipulated optical slant by changing the viewing perspective while maintaining the geographical slant. They did not find an effect of perspective, demonstrating that observers might not perceive optical slant.

For a slanted surface rotating to yield structure-from-motion, optical slant would constantly change. If observers truly perceived optical slant, slant judgments should be highly variable in this circumstance. This was not what I found and instead, slant judgments were predicted well by geographical slant. In the context of SFM, the perception of slant in terms of geographical slant means that the momentary direction of the slant cannot be determined in terms of the direction of maximum increasing depth (e.g. Stevens, 1983a, 1983b). Because of this, observers may be somewhat imprecise in determining the direction of slant. If so, then observers might rely on the direction of major axes of symmetry of a surface contour shape as a constraint used to estimate the slant direction. Indeed, I found that slant judgments were relatively poor in the absence of symmetry and good with symmetry, and that a change in the direction of the symmetry axis impacted slant judgments as predicted by this hypothesis. Finally, most studies on optical slant and tilt perception suggested that tilt is perceived reliably and accurately whereas slant is not (Stevens, 1983b; Koenderink & van Doorn, 1995; Koenderink, van Doorn, & Kappers, 1992, 1994, 1995; Norman, Todd, Phillips, 1995; Todd, Koenderink, van Doorn, & Kappers, 1996). Now, with a geographical frame of reference and continuous perspective change, slant is perceived accurately but perception of tilt (or the direction of slant) is a function of the symmetry of the surface contour.

In addition, local slant has also been considered as the basic component of 3D shape (e.g. Koffka, 1935; Gibson, 1950; Beck & Gibson, 1952; Wallach & Moore, 1962; Kaiser, 1967;

Hoffman & Richards, 1984; Todd, 2004; Sakata, Tsutsui & Taira, 2005; Welchman, Deubelius, Conrad, Bültoff & Kourtzi, 2005). However, results from this dissertation suggested that this might not be true for several reasons. In addition to the fact that participants did not perceive local slant, slant judgment failed to improve with large rotation in the absence of additional 3D structure on the planar surface. This essentially means that one has to make planar surfaces to be more 3D-shape-like to allow accurate slant estimates. Overall, evidence presented in this dissertation suggested that slant should be considered as a characteristic of 3D shapes, along with other characteristics, such as aspect ratios. For bootstrapping 3D metric structures, the recovered scaling factor should be applied to the 3D object as a whole. In Experiment 4 where I used asymmetric pentagonal objects, slant judgments failed to improve due to the ambiguity of the direction of slant. However, as I mentioned, this did not mean that participants failed to recover the correct scaling factor with large rotation. In other words, judgment of other characteristics of the object, such as aspect ratio, could still be accurate. This way of thinking could be further explored in subsequent experiments.

Finally, I should note that this understanding of perceived surface orientation is consistent with the phenomenological experience of surroundings that remain stable when we move about changing our perspective on them so as to provide a basis for reliable and effective actions, locomotion through the surroundings, moving to lean or sit on a surface, or reaching out to grab an object. We intuitively move relative to perceived surfaces in order to gain a better appreciation of what they are and how they are configured in 3D. The bootstrap process, with its application to improve the perception of the metric 3D structure of surrounding surfaces is consistent with this common experience and practice. Altogether, perception of geographical slant, in contrast to optical slant, makes sense.

References

Allen, B., Haun, A. M., Hanley, T., Green, C. S., & Rokers, B. (2015). The optimal combination of the binocular cues to 3d motion. *Investigative Ophthalmology & Visual Science*, *56*(12), 7589-7596.

Bhalla, M., & Proffitt, D. R. (1999). Visual–motor recalibration in geographical slant perception. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 1076.

Beck, J., & Gibson, J.J. (1955). The relation of apparent shape to apparent slant in the perception of objects. *Journal of Experimental Psychology* 50(2), 125-133.

Berkeley, G. (1709/2002). *An Essay Towards A New Theory of Vision.* Aaron Rhames.

Bingham, G. P., & Muchisky, M. M. (1993a). Center of mass perception and inertial frames of reference. *Perception & Psychophysics*, *54*(5), 617-632.

Bingham, G. P., & Muchisky, M. M. (1993b). Center of mass perception: Perturbation of symmetry. *Perception & Psychophysics*, *54*(5), 633-639.

Birkhoff, G. & Mac Lane, S. (1996). Affine geometry. In *A Survey of Modern Algebra, 5th ed.* New York: Macmillan, 268-275.

Brainard, D. H. (1997) The Psychophysics Toolbox, *Spatial Vision 10*:433-436.

Cherry, O. C., & Bingham, G. P. (2018). Searching for invariance: Geographical and optical slant. *Vision Research*, *149*, 30-39.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42-45.

Cumming, B.G., & Parker, A.J. (1994). Binocular mechanisms for detecting motion-in-depth. *Vision Research*, 34(4), 483-495.

Dick, M., & Hochstein, S. (1989). Visual orientation estimation. *Perception & Psychophysics*, *46*(3), 227-234.

Domini, F., & Caudek, C. (2013). Perception and action without veridical metric reconstruction: an affine approach. In S. Dickinson, & Z. Pizlo (Eds.) *Shape Perception in Human and Computer Vision* (pp. 285-298). Springer London.

Durgin, F. H., & Li, Z. (2011). The perception of 2D orientation is categorically biased. *Journal of Vision*, *11*(8), 13-13.

Durgin, F. H., Li, Z., & Hajnal, A. (2010). Slant perception in near space is categorically biased: Evidence for a vertical tendency. *Attention, Perception, & Psychophysics*, *72*(7), 1875-1889.

Firestone, C. (2013). How "paternalistic" is spatial perception? Why wearing a heavy backpack doesn't—and couldn't—make hills look steeper. *Perspectives on Psychological Science*, *8*(4), 455-473.

Flock, H. R. (1964). Some conditions sufficient for accurate monocular perceptions of moving surface slants. *Journal of Experimental Psychology*, 67, 560-572.

Gibson, J.J. (1950). The perception of visual surfaces. *The American Journal of Psychology*, 63(3), 367-384.

Gibson JJ (1986) *The Ecological Approach to Visual Perception*. Psychology Press, New York.

Gibson, J.J., & Cornsweet, J. (1952). The perceived slant of visual surfaces – optical and geographical. *Journal of Experimental Psychology,* 44(1), 11-15.

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*, *4*(12), 967-992.

Hoffman, D.D., & Richards, W.A. (1984). Parts of recognition. *Cognition*, 18, 65-96.

Huang, T. S., & Netravali, A. N. (2002). Motion and structure from feature correspondences: A review. *Proceedings of IEEE*, 82(2), 252-268.

Johnston, E.B. (1991). Systematic distortions of shape from stereopsis. *Vision Research*, *31*(7-8), 1351-1360.

Julesz, B. (1971). *Foundations of Cyclopean Perception.* Chicago: The University of Chicago Press.

Kaiser, P.K. (1967). Perceived shape and its dependency on perceived slant. *Journal of Experimental Psychology*, 75(3), 325-353.

Klein, F. (1893). A comparative review of recent researches in geometry. *Bulletin of the American Mathematical Society*, 2(1), 215-249.

Kleiner M., Brainard D., Pelli D. (2007). What's new in Psychtoolbox-3?. *Perception*, 36, ECVP Abstract Supplement.

Knill, D. C. (2007). Learning Bayesian priors for depth perception. *Journal of Vision*, *7*(8), 13-13.

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant?. *Vision Research*, *43*(24), 2539-2558.

Koenderink, J. J., & Van Doorn, A. J. (1991). Affine structure from motion. *JOSA A*, *8*(2), 377-385.

Koenderink, J. J., & Van Doorn, A. J. (1995). Relief: Pictorial and otherwise. *Image and Vision Computing*, *13*(5), 321-334.

Koenderink, J. J., Van Doorn, A. J., & Kappers, A. M. (1992). Surface perception in pictures. *Perception & Psychophysics*, *52*(5), 487-496.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. (1994). On so-called paradoxical monocular stereoscopy. *Perception*, *23*(5), 583-594.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. (1995). Depth relief. *Perception*, *24*(1), 115-126.

Koffka, K. (1935). *Principles of Gestalt Psychology. New York, NY; Harcourt Brace.*

Lappin, J. S., & Fuqua, M. A. (1983). Accurate visual measurement of three-dimensional moving patterns. *Science*, *221*(4609), 480-482.

Lee, Y. L., & Bingham, G. P. (2010). Large perspective changes yield perception of metric shape that allows accurate feedforward reaches-to-grasp and it persists after the optic flow has stopped!. *Experimental Brain Research*, *204*(4), 559-573.

Lee, Y. L., Crabtree, C. E., Norman, J. F., & Bingham, G. P. (2008). Poor shape perception is the reason reaches-to-grasp are visually guided online. *Perception & Psychophysics*, *70*(6), 1032-1046.

Lee, Y. L., Lind, M., Bingham, N., & Bingham, G. P. (2012). Object recognition using metric shape. *Vision Research*, *69*, 23-31.

Li, Y., Sawada, T., Shi, Y., Kwon, T., & Pizlo, Z. (2011). A Bayesian model of binocular perception of 3D mirror symmetrical polyhedra. *Journal of Vision*, *11*(4), 1-20.

Lind, M. (1996). Perceiving motion and rigid structure from optic flow: A combined weak-perspective and polar-perspective approach. *Perception & Psychophysics*, *58*(7), 1085-1102.

Lind, M., Lee, Y. L., Mazanowski, J., Kountouriotis, G. K., & Bingham, G. P. (2014). Affine operations plus symmetry yield perception of metric shape with large perspective changes (≥ 45°): Data and model. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 83.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press. *Cambridge, Massachusetts*.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61-64.

Nefs, H.T., O'Hare, L., & Harris, J.M. (2010). Two independent mechanisms for motion-in-depth perception: Evidence from individual differences. *Frontiers in Psychology*, 1.

Nistér, D. (2005). Preemptive RANSAC for live structure and motion estimation. *Machine Vision and Applications*, *16*(5), 321-329.

Norman, J. F., Crabtree, C. E., Bartholomew, A. N., & Ferrell, E. L. (2009). Aging and the perception of slant from optical texture, motion parallax, and binocular disparity. *Perception & Psychophysics*, *71*(1), 116-130.

Norman, J. F., & Lappin, J. S. (1992). The detection of surface curvatures defined by optical motion. *Perception & Psychophysics*, *51*(4), 386-396.

Norman, J. F., & Todd, J. T. (1993). The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception & Psychophysics*, *53*(3), 279-291.

Norman, J. F., & Todd, J. T. (1996). The discriminability of local surface structure. *Perception, 25(4)*, 381-398.

Norman, J. F., Todd, J. T., & Phillips, F. (1995). The perception of surface orientation from multiple sources of optical information. *Perception & Psychophysics*, *57*(5), 629-636.

Norman, J.F., Todd, J.T., Perotti, V.J., & Tittle, J.S. (1996). The visual perception of three-dimensional length. *Journal of Experimental Psychology: Human Perception and Performance*, 22(1), 173-186.

Norman, J. F., Todd, J. T., Norman, H. F., Clayton, A. M., & McBride, T. R. (2006). Visual discrimination of local surface structure: Slant, tilt, and curvedness. *Vision Research*, *46*(6-7), 1057-1069.

Palmer, S. E. (1985). The role of symmetry in shape perception. *Acta Psychologica*, *59*(1), 67-90.

Pelli, D. G. (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision 10*:437-442.

Perotti, V. J., Todd, J. T., Lappin, J. S., & Phillips, F. (1998). The perception of surface curvature from optical motion. *Perception & Psychophysics*, *60*(3), 377-388.

Pizlo, Z. (2010). *3D Shape: Its Unique Place in Visual Perception*. Cambridge, MA: MIT Press.

Pizlo, Z., Sawada, T., Li, Y., Kropatsch, W. G., & Steinman, R. M. (2010). New approach to the perception of 3D shape based on veridicality, complexity, symmetry and volume. *Vision research*, *50*(1), 1-11.

Proffitt, D. R., Creem, S. H., & Zosh, W. D. (2001). Seeing mountains in mole hills: Geographical-slant perception. *Psychological Science*, *12*(5), 418-423.

Purdy, J., & Gibson, E. J. (1955). Distance judgment by the method of fractionation. *Journal of Experimental Psychology*, *50*(6), 374.

Sakata, H., Tsutsui, K.-I., & Taira, M. (2005). Toward an understanding of the neural processing for 3D shape perception. *Neuropsychologia, 43,* 151-161.

Saunders, J. A., & Chen, Z. (2015). Perceptual biases and cue weighting in perception of 3D slant from texture and stereo information. *Journal of Vision*, *15*(2): 14, 1-24.

Saunders, J. A., & Knill, D. C. (2001). Perception of 3D surface orientation from skew symmetry. *Vision Research*, *41*(24), 3163-3183.

Sawada, T., & Pizlo, Z. (2008). Detection of skewed symmetry. *Journal of Vision*, *8*(5): 14, 1-18.

Sedgwick, H.A. & Levy, S. (1985). Environment-centered and viewer-centered perception of surface orientation. *Computer Vision, Graphics, and Image Processing, 31,* 248-260.

Shapiro, L.S., Zisserman, A., & Brady, M. (1995). 3D motion recovery via affine epipolar geometry. *International Journal of Computer Vision*, *16*(2), 147-182.

Shioiri, S., Saisho, H., & Yaguchi, H. (2000). Motion in depth based on inter-ocular velocity differences. *Vision Research*, 40, 2565-2572.

Stevens, K.A. (1983a). Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics, 46*, 183-195.

Stevens, K. A. (1983b). Surface tilt (the direction of slant): A neglected psychophysical variable. *Perception & Psychophysics*, *33*(3), 241-250.

Tittle, J.S., & Braunstein, M.L. (1993). Recovery of 3-D shape from binocular disparity and structure from motion. *Perception & Psychophysics*, *54*(2), 157-169.

Tittle, J. S., Todd, J. T., Perotti, V. J., & Norman, J. F. (1995). Systematic distortion of perceived three-dimensional structure from motion and binocular stereopsis. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 663.

Todd, J.T. (2004). The visual perception of 3D shape. *Trends in Cognitive Sciences*, 8(3), 115-121.

Todd, J.T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception and Psychophysics,* 48(5), 419-430.

Todd, J.T., Christensen, J.C., & Guckes, K.M. (2010). Are discrimination thresholds a valid measure of variance for judgments of slant from texture?. *Journal of Vision*, 10(2), 1-18.

Todd, J. T., Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. (1996). Effects of changing viewing conditions on the perceived structure of smoothly curved surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(3), 695.

Todd, J. T., & Norman, J. F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception & Psychophysics*, *50*(6), 509-523.

Todd, J. T., Oomes, A. H., Koenderink, J. J., & Kappers, A. M. (2001). On the affine structure of perceptual space. *Psychological Science*, *12*(3), 191-196.

Todd, J.T., & Perotti, V.J. (1999). The visual perception of surface orientation from optical information. *Perception & Psychophysics,* 61(8), 1577-1589.

Todd, J. T., Thaler, L., & Dijkstra, T. M. (2005). The effects of field of view on the perception of 3D slant from texture. *Vision Research*, *45*(12), 1501-1517.

Todd, J.T., Tittle, J.S., & Norman, J.F. (1995). Distortions of three-dimensional space in the perceptual analysis of motion and stereo. *Perception*, 24(1), 75-86.

Ullman, S. (1983). Recent computational studies in the interpretation of structure from motion. *Human and Machine Vision*, 459-480.

van Ee, R., Adams, W. J., & Mamassian, P. (2003). Bayesian modeling of cue interaction: bistability in stereoscopic slant perception. *JOSA A*, *20*(7), 1398-1406.

Wagner, M. (1985). The metric of visual space. *Perception & Psychophysics*, *38*(6), 483-495.

Wagner, M. (2008). Comparing the psychophysical and geometric characteristics of spatial perception and cognitive maps. *Cognitive Studies*, 15(1), 6-21.

Wallach, H., & Moore, M. E. (1962). The role of slant in the perception of shape. *American Journal of Psychology*, *75*, 285–293.

Wang, X.M., Lind, M., & Bingham, G.P. (2018). Large continuous perspective change with noncoplanar points enables accurate slant perception. *Journal of Experimental Psychology: Human Perception and Performance*, 44(10), 1508-1522.

Welchman, A.E., Deubelius, A., Conrad, V., Bültoff, H.H. & Kourtzi, Z. (2005). 3D shape perception from combined depth cues in human visual cortex. *Nature Neuroscience, 8,* 820-827.

Zisserman, A., Beardsley, P. A., & Reid, I. D. (1995). Metric calibration of a stereo rig. In *Proceedings IEEE Workshop on Representation of Visual Scenes (In Conjunction with ICCV'95)* (pp. 93-100). IEEE.

Curriculum Vitae

Xiaoye 'Michael' Wang
Department of Psychological and Brain Sciences
Indiana University
Bloomington, IN 47408
E-mail: wang492@indiana.edu

EDUCATION

B.A. Major: Psychology; minor: Philosophy, *cum laude,* May 2013
Denison University
Granville, OH 43023

Ph.D. Psychology, and Cognitive Science, May 2019
Indiana University
Bloomington, IN 47405

PUBLISHED PAPERS

Snapp-Childs, W., **Wang, X. M**., & Bingham, G. P. (2016). Progressive reduction versus fixed level of support during training: When less is less. *Human Movement Science*, *45*, 172-181.

Pan, J.S., Li, J., Chen, Z., Mangiaracina, E.A., Connell, C.S., Wu, H., **Wang, X.M**., Bingham, G.P., & Hassan, S.E. (2017). Motion-Generated Optical Information Allows Event Perception despite Blurry Vision in AMD and Amblyopic Patients. *Journal of Vision,* 17(12):13, 1-16.

Wang, X.M., Lind, M., & Bingham, G. P. (2018). Large continuous perspective change with non-coplanar points enables accurate slant perception. *Journal of Experimental Psychology: Human Perception and Performance*. *44*(10), 1508-1522.

Wang, X.M., & Bingham, G.P. (2019). Change in effectivity yields recalibration of affordance geometry to preserve functional dynamics. *Experimental Brain Research*, 237(3), 817-827.

Wu, H.*, **Wang, X.M.***, & Pan, J.S. (2019). Perceiving blurry scenes with translational optic flow, rotational optic flow or combined optic flow. *Vision Research*, 158, 49-57.
*these authors contributed equally to this work.

MANUSCRIPT IN PREPARATION/UNDER REVIEW

Wang, X.M., Lind, M., & Bingham, G.P. (under review). Symmetry mediates the bootstrapping of 3D relief slant to metric slant. *Vision Research*.

Wang, X.M., Lind, M., & Bingham, G.P. (under review). Bootstrapping a better slant: A stratified process for recovering 3D metric slant using a bootstrap process. *Vision Research.*

INVITED TALKS

Bingham, G.P., Lind, M., Fath, A.J., **Wang, X.M.** (2015). Perceiving Affine and Metric 3D Surfaces Using Multiple Motion Systems. An invited paper presented at the Department of Psychology, University of Connecticut, Storrs, CT, May 28th.

PUBLISHED ABSTRACTS

Wang, X.M., Fath, A.J., Lind, M., & Bingham, G. (2015). A New Slant on" Two Eyes Are Better Than One": Large Continuous Perspective Changes (≥ 45°) Allow Metric Slant Perception Using Cyclopean (or Stereo-) Motion. *Journal of Vision*, *15*(12), 724-724.

Wang, X.M., Fath, A.J., Snapp-Childs, W., Lind, M., & Bingham, G.P. (2016). Inhomogeneity of perceived slants with different motion-based visual information. *Journal of Vision*, 16(12), 654.

Wang, X.M., Lind, M., & Bingham, G.P. (2018). Modeling 3D Slant Perception with Motion-Based Visual Information. *Journal of Vision*, 18(10), 129.

POSTERS

Wang, X.M., Pan, J.S., Bingham, B.P. (2014, June). Global optical flow aids scene perception under conditions simulating low vision. Poster session presented at the International Society for Ecological Psychology North American Meeting, Miami, OH, USA.

Wang, X.M., Pan, J.S., Li, J., Chen, Z., Mangiaracina, E.A., Connell, C.S., Wu, H., Bingham, G.P., Hassan, S.E. (2017 June). Motion-generated optical information allows event perception despite blurry vision in AMD and amblyopic patients. Poster session presented at the Workshop on Ecocentric Vision: From Science to Real-World Applications, Bloomington, IN, USA.

MENTORSHIP/STUDENT SUPERVISION

Undergraduate Research Assistants:

Bradley Perfitt (Psychology), Fall 2014 – Spring 2015
Senior Thesis: “The effect of unknown rotation on embodied memory”.

Bolin Liao (Psychology), Fall 2014 – Spring 2015

Senior Thesis: “Stereo visual information along with motion information is necessary for accurate slant perception”.

Alexander Campell (Cognitive Science), Fall 2015 – Spring 2016
Senior Thesis: “Investigation of slant-perceiving agent”.

Kelli Steinbuck (Psychology), Fall 2016
Final Project: “The effects of 3-Dimensional Corrugation on Slant Perception Using Stereomotion and Monocular Motion”

Christian Willenbacher (Cognitive Science), Spring 2019
Final Project: “The specification of the direction of geographical slant”

TEACHING/TEACHING ASSISTANT EXPERIENCES

Associate Instructor:

PSY-P 211 (Methods of Experimental Psychology), Indiana University
Term(s): Fall 2015

Teaching Assistant:

PSY-P 330 (Perception and Action), Indiana University
Term(s)Fall 2013, Spring 2014, Fall 2014, Spring 2015, Spring 2016, Fall 2016, Spring 2017, Fall 2017, Spring 2018, Fall 2018.

PSY-K 300 (Statistical Techniques), Indiana University
Term(s): Summer 2014.

PSY-P 324 (Abnormal Psychology), Indiana University
Term(s)Summer 2015, Summer 2016, Summer 2017, Summer 2018.

PSY-P 459/657 (History and Systems of Psychology), Indiana University
Term(s): Spring 2019.

AWARDS AND MERITS

Indiana University College of Arts and Sciences Spring Travel Award, 2018, College of Arts and Sciences, Indiana University, Bloomington, IN, USA.

Research Fellowship, 2018, Cognitive Science Program, Indiana University, Bloomington, IN, USA.

Research Fellowship, 2016, Cognitive Science Program, Indiana University, Bloomington, IN, USA.

Fellow, 2013, Psychology Department, Denison University, Granville, Ohio.

R. Stanley and Janet O. Laing Scholarship, 2012 – 2013 Denison University, Granville, Ohio.

Dean's List, 2010, 2011, 2012, 2013 Denison University, Granville, Ohio.

Ronald S. Beard Endowed Scholarship, 2009 – 2013 Denison University, Granville, Ohio.

PROFESSIONAL ORGANIZATIONS AND AFFILIATIONS

International Society for Ecological Psychology
International Vision Science Society
Psi Chi, International Honor Society in Psychology
Omicron Delta Epsilon, International Honor Society for Economics